ELECTRICAL COMMUNICATION

ITT

VOLUME 38 • NUMBER 3 • 1963

ELECTRICAL COMMUNICATION

Technical Journal Published Quarterly by

INTERNATIONAL TELEPHONE and TELEGRAPH CORPORATION

320 Park Avenue, New York 22, New York

President: H. S. Geneen

Secretary: J. J. Navin

CONTENTS

Volume 38 **1963** **Number 3**

EDITOR, H. P. Westman

ASSISTANT EDITOR, M. Karsh

Subscription: $2.00 per year

50¢ per copy

This Issue in Brief

Trends in Component Developments—Much of the advancement in electronic systems has stemmed from progress in the development of essential components. For entertainment equipment, stress on price stimulated mechanization. The military market added severe environmental conditions and a lower failure rate. Later, special telecommunication and electronic systems demanded failure rates even lower than the military, closer tolerances, and lives of the order of tens of years. Presently, these and the military requirements have tended to coalesce.

Reliability can no longer be the process of selecting the best from commercial production but now requires special design and manufacture in clean areas by workers of high integrity who will not pass an inferior unit. Improved reliability has raised problems in life testing and quality control.

Interconnections between components and interfaces of different materials within components are additional sources of trouble. Two possible solutions are the use of a single chip of semiconductor material having both active and passive elements built into it and the use of a substrate of glass with passive elements deposited on it. These integrated circuits will be economical only if they are produced in large quantities and the difficulties in standardizing components indicate that it will not be easy to obtain acceptance of standardized circuits.

The new components with their demands for high purity of materials and close control of manufacturing techniques have brought the physicist and chemist into the components industry. Mechanization is no longer a cost-reduction tool but an essential means of attaining uniformity and reliability. Dust particles now are significant in size compared with the product, and dust-free regions, both temperature and humidity controlled, are found within the factory.

The epitaxial process of producing planar transistors has not only increased the operating range and power of these active components but has greatly improved life over other encapsulation methods.

More-accurate machining techniques, parametric amplifiers, lasers, vacuum deposition, molecular excitation, high-resolution cathode-ray tubes, and other techniques and components are continually extending the useful upper frequency limits of existing systems.

High-Reliability Testing and Assurance for Electronic Components—The reliability of an electronic component is the probability that it will perform a specific function for a specified length of time under defined environmental conditions. By nature, it is a quantity that is predicted rather than measured.

The determination of average levels of reliability does not present any exceptional difficulties. But as reliability is tending toward the increasingly higher values that are demanded, theoretical, practical, and economic difficulties are arising. To overcome them, various means such as accelerated tests, which also are debatable, are available. The problem becomes clearer and simpler if a separation is made between random unforeseeable failures and failures due to ageing, which are easier to predict and are of prime importance in long-term applications.

Finally, a broad organization of reliability, based on cooperation, has to be envisaged.

Microelectronic Devices—Three types of microelectronic devices are discussed. Integrated circuits consist of a number of interconnected components that perform a specific function on a common mount. Thin-film circuits use resistors and capacitors that are evaporated or sputtered on a common mount and the active elements are added separately. Semiconductor circuits have both active and passive elements processed on a semiconductor substrate by transistor techniques.

Integrated tunnel-diode circuits show spectacular improvements over lumped components.

Close matching of characteristics and low series inductance of leads and elements are important factors and may permit operation at speeds of several hundred megacycles per second. Smallness and compact arrangement of elements reduce propagation time over the signal path. Examples of logic, memory, and counting units are described but the techniques are not limited to tunnel diodes.

Frequently, transistors are used in matched sets and by mounting them in intimate thermal contact the matching will be improved over the full operating temperature range. The mounting in a single header of the two transistors that make up a Darlington amplifier increases reliability by reducing the number of connections between dissimilar materials as well as requiring less space and lower encapsulating costs.

Means of evaporating active devices on to the thin-film substrate of glass or ceramic to which the passive elements are applied are still under development. The Miniflake transistor is readily attached to the substrate, though not part of it.

In semiconductor circuits, the transistor diffusion techniques permit active and passive elements to be incorporated in the substrate. The semiconductor substrate may be used as a common electrode or alternatively isolation may be obtained by using a reverse-biased junction. Passive components may also be produced on the semiconductor circuit by thin-film techniques.

Semiconductor circuits should provide high reliability as all components are processed simultaneously of very pure materials under carefully controlled conditions. Packing density is limited by heat dissipation and low-power devices can be made very small. Tooling costs being high, production runs must be large. The active elements are most suited to digital operation.

Thin-film techniques have lower tooling costs and are simpler to design than semiconductor units. The active elements being added, the thin-film circuits may be better suited for linear operation. Semiconductor and thin-film devices are not necessarily competitive but have their own fields of usefulness.

Silring Mounting of Silicon Power Rectifiers—Based on the techniques of stacking selenium rectifiers, silicon units are mounted in doughnut-shaped structures that are alternately stacked with cooling plates on an insulated rod. The rectifier units serve as spacers between the cooling plates.

The rectifier element is soldered to a heavy copper ring and fits into an eccentric opening in a ceramic spacer. A second copper ring provides for the other connection to the rectifier and both plates and the ceramic insulator are soldered together to provide hermetic sealing.

The rectifier cells, cooling plates, and connection tabs may be mounted in any electrical arrangement on an insulated bolt passing through their centers. At an ambient temperature of 50 degrees centigrade, ratings between 2.5 amperes for the Silring unit alone to 15 amperes with cooling plates 125 millimeters (4.9 inches) square by 2 millimeters (0.08 inch) thick are obtained. Parallel operation of rectifiers of the same rating, but without sorting of units, permits direct-current outputs of 280 amperes or twice this current for forced cooling with an air speed of 6 meters (19.7 feet) per second. The relative sizes of stacks for approximately the same rating are 100 percent for selenium, 8 percent for conventional silicon mounting, and 4 percent for Silring.

Cause and Prevention of High Reverse Currents in Large-Area High-Voltage Diffused-Silicon Rectifiers—It is shown that the major factor responsible for "soft" reverse characteristics in silicon diodes is the presence of impurities, other than the elements of Groups III and V, distributed uniformly throughout the silicon. The impurities are present in both zone- and

crucible-grown silicon with inevitable further contamination during device processing.

An expression is derived relating the reverse current of such soft diodes to voltage and temperature, and a phenomenological explanation of the processes underlying this relationship is presented.

Two methods of removing the impurities from the vicinity of the junction are described. The first method, in which an acidic oxide is used as a getter, was first described by Goetzberger and Shockley. The second, a physical gettering process, has not been used previously to improve the quality of soft diodes.

The mechanism of these processes has been studied with a view to defining the optimum conditions under which they proceed. It is suggested that these techniques provide a means of stabilizing device yield at a high level, as well as the possibility of producing new devices by complete removal and subsequent controlled addition of the impurities.

Defects in Vapour-Grown Silicon—Single-crystal silicon may be deposited from a vapour on a single-crystal slice of the same material. Diffusion of impurities is reduced if the substrate is maintained at a temperature substantially lower than that of the melting point and, by suitable doping of the vapour, p-n junctions can be grown with predetermined impurity profiles.

Crystalline defects are reduced by annealing and cleaning the mechanically polished substrate on which the epitaxial layer is deposited, indicating that imperfections or surface impurities encourage such defects. Dislocation densities are not increased in deposited layers over those in the substrate with proper annealing and cleanliness.

Polycrystallinity and twinning are encouraged by contamination and by growth at too low a temperature. Stacking faults appear to originate at either a physical or a chemical imperfection but further study is needed of the nucleation process.

With adequate precautions, vapour-grown layers of silicon have defect concentrations similar to those of the substrate on which they are grown.

Silicon Epitaxial Planar Transistors—Two major shortcomings of normal diffused mesa transistors are high saturation voltage V_{CES} and long switching time, both being closely dependent on current levels and the latter also being influenced by the circuit arrangement. Noise and low current gain at low currents are additional problems. Epitaxial technology and gold diffusion have overcome the two major and the planar technology has overcome the other two shortcomings. They have also reduced manufacturing tolerances of characteristics, which are also retained reliably during life.

Epitaxial layers of silicon suitably doped to control resistivity have the same crystalline orientation as the substrate on which they are grown and may be extremely thin to reduce the series electric and thermal resistances. Epitaxy produces transistors having low saturation resistances that are independent of temperature, increased linearity of characteristics, and reduced minority-carrier storage time.

By diffusing gold into the base and collector regions, the storage time in switching circuits has been reduced from 100 to 10 nanoseconds in a typical low-level epitaxial transistor.

In the planar technique, an oxide layer less than a micron thick is grown on a clean silicon surface. A window is etched in this silicon dioxide and, exposed to a high-temperature boron-rich atmosphere, boron is diffused, producing a p-type layer underneath a new oxide layer formed over the window. Similarly another window is etched and an n-type layer is formed beneath another oxide coating by phosphorus. Thus, each junction is completely sealed by an oxide layer as it is formed, permanently preventing contamination. Leads are connected through small etched windows.

The epitaxial and planar techniques as well as

gold doping are combined, making the silicon epitaxial planar transistor. These are probably the most versatile and reliable transistors available. The production techniques are inexpensive and over a thousand units may be processed simultaneously. Tests at each stage of manufacture avoid further processing of rejects.

Characteristics of a number of typical planar and planar epitaxial silicon transistors are given. Epitaxial and non-epitaxial units are compared for direct-current, alternating-current, and switching characteristics.

Mixer Diodes for 6 Gigacycles Per Second—Two formed germanium diodes, both operating on nonlinear-resistance properties, were developed for down-converter and up-converter service between 70 megacycles per second and 6 gigacycles per second.

In the low-level down-converter, noise sets the sensitivity limit of the receiver. Noise, conversion efficiency, and output impedance at the intermediate frequency are inversely related to the local-oscillator drive, which determines the voltage swing and thus the extent of nonlinearity of the operating range.

Although silicon diodes are simpler to make, requiring only a pressure contact between the whisker and the polycrystalline semiconductor, in contrast to a formed contact in a germanium diode in which the whisker is welded by an electric current to a monocrystalline semiconductor, the germanium has lower series resistance and lower intermediate-frequency impedance, the latter being an important element for this design. The theoretical superiorities of gallium arsenide have not been obtained in production with assurance.

As the forming current and resulting heat is increased, the area of the junction also increases. The series resistance of the inactive part of the semiconductor varies as the reciprocal of the junction radius and the capacitance in shunt of the active region varies as the square of the junction radius. By putting two whiskers in parallel, the changes in series resistance and shunt capacitance are proportionate but the intermediate-frequency impedance is halved. Curves are given for noise, conversion loss, and intermediate-frequency impedance of the final design of diode.

An up-converter using nonlinear capacitance is more efficient then the nonlinear-resistance model. However, the series resistance of the latter can be used effectively as a means of limiting output amplitude. Also by making a large junction, its shunt capacitance is increased and harmonics are short-circuited. The *p-i-n* type, separating the *p* and *n* regions by an intrinsic layer, provides a larger junction the capacitance of which is invariant and not significantly sensitive to voltage.

If *n*-type germanium is heated sufficiently, it will pass through the intrinsic state and become *p* type. By using a high-resistivity *n*-type germanium, a *p*-type-impurity whisker, and the proper forming technique, a heavily doped *p* region occurs immediately under the whisker, a balanced intrinsic region of high resistivity is beyond it, and the original *n*-type germanium is beyond that.

The impedance change from $+1$ to -1.6 volts is predominantly resistive, the conversion loss at saturation drive is only 7.5 decibels, and the compression ratio is 5 : 1 for an input change of 5.6 decibels. All interfering harmonics were at least -117 decibels relative to the output signal.

Vacuum Tubes for Submerged Repeaters—The cost of an interruption in service and of replacing a submarine repeater on a multichannel telephone system is such as to justify almost any expense to insure reliability for a life of, say, 20 years.

The vacuum tubes, which are the essential active elements in the repeaters, must have high transconductance combined with low interelectrode capacitance to obtain the largest frequency bandwidth with the minimum number of repeaters. They must operate at low voltages and currents to impose the least demands on the

power supply that operates in series all the repeaters in a system. The design must balance these often-conflicting requirements.

The required reliability can be attained only by perfection in every stage of manufacture including the highest standards of cleanliness. The entire process from raw materials to life testing is done in a fully restricted area completely isolated from the rest of the plant.

Characteristics and some historical information are given on the following types: *5A/153, 5A/161, 5A/162, 5A/181, 5A/182, LS882,* and *VX7158*. Transconductance changes over many thousands of hours of operation indicate extremely small variations. About 18 million tube-hours have been accumulated in submarine repeaters from 1955 to 1962 and their reliability well justifies the faith placed in the design and production of these vacuum tubes.

Ceramic-Insulated Vacuum Tubes for Very-High-Frequency Industrial Heating—The production of larger powers at higher frequencies will expand the uses of dielectric heating in industry and has prompted the development of four new triodes producing up to 12 kilowatts at 220 megacycles per second and 40 kilowatts at 100 megacycles per second.

Tapered coaxial elements permit smaller structures having lower inductances than corresponding conventional pin-type tubes and give better performance at the higher frequencies. A high-alumina ceramic insulation between adjacent tube elements has lower dielectric losses than glass and its use also has the advantage of a resultant low radio-frequency loss at the seals. These ceramic seals may operate at 220 to 250 degrees centigrade against 180 degrees for glass-Kovar seals.

As exact matching of the coefficients of expansion of both ceramic and metal cannot be obtained from brazing to room temperature, the metal is selected to have the larger coefficient and it is brazed to the outside of the ceramic rings, thus operating the ceramic in compression.

The filament is operated at a relatively low temperature and with uniform optimum carburization will give a minimum life of 5000 hours with an unregulated filament voltage that may vary ± 5 per cent and will operate for short times at 90 per cent of rated voltage.

The grid design is particularly critical because if the tube is operated without a load, the grid dissipation can be several times the normal loaded value. A much-larger safety factor must be designed into the grid than would be required for a conventional radio transmitting tube.

Air, water, and vapour cooling are provided for in the several designs. The trend in medium-size equipments is to the use of a water-cooling coil hard soldered to the anode wall.

Travelling-Wave Tubes for 6-Gigacycle-Per-Second Radio Links—Two new travelling-wave tubes were developed for use as output amplifiers in 6-to-7-gigacycle-per-second radio links. One tube with a synchronous helix voltage of 3 kilovolts and a beam current of 40 milliamperes gives a saturated output power of over 10 watts. The other gives about 50-percent more power for use in the highest-capacity radio links and operates at 3.5 kilovolts and 50 milliamperes.

The tubes include the same cathode and electron guns used in earlier designs of tubes operating at 4 gigacycles per second as these have demonstrated an average life of over 20-000 hours. A novel design of mount, with a high degree of radio-frequency screening, incorporates the periodic permanent magnets used for focusing the electron beam, the waveguide feeds, and a natural-convection cooler. The tubes can be easily replaced in the mounts under field conditions and will give very good focusing.

Results of measurements of noise factor (around 26 decibels), modulation noise, gain (37 to 40 decibels), power output, and other characteristics are given.

Effect on an Electron Beam of Variations in Periodic Permanent-Magnet Focusing Systems —Theoretical solutions for electron beam focusing in uniform periodic permanent-magnet systems exist, but in practical cases it is important to know the effect of the variations and misalignments that inevitably occur.

In the paper the results of computer calculations are given for the behavior of an electron beam subject to random variations of magnetic field and half-period length. These show the magnetic field to be the more-critical parameter. The effect of a change in magnetic field has also been calculated, and it is shown that a gradual stepping-up of field will maintain a well-focused beam. Also a moderate increase in period length, such as may be necessary to accommodate a waveguide, can be compensated for by a slight change in magnetic field. If larger changes of period are necessary, the pole-pieces can be incorporated in the waveguide itself without adversely affecting the matching; details of this are given.

Another important variable will be the displacement of the various elements of the periodic focusing system from the desired axis. The resultant beam displacement is derived analytically leading to a value, for practical purposes, of $2sn^{\frac{1}{2}} \left|\sin (2\alpha)^{\frac{1}{2}}\right|$ where n is the number of elements, and s the standard deviation of the displacement of the elements.

Recent Achievements

Smithsonian Institution Receives Microwave Relics—Over thirty years ago, microwave communication was first publicly demonstrated by transmitting across the English Channel from Escalles, France, to Saint Margaret's Bay in England. Both telegraphy and telephony were tested. Operating at 1.7 gigacycles per second, a half-watt transmitter energized a half-wave dipole antenna mounted in a 10-foot (3-meter) reflector.

The Smithsonian Institution in Washington, District of Columbia, has accepted for display the original transmitter section with vacuum tubes and the antenna dipole assembly together with a replica of the 10-foot reflector used in the demonstration. The presentation, shown in Figure 1, was made by Mr. Andre G. Clavier, who led the engineering team that developed the equipment. The International Telephone and Telegraph Laboratories at Hendon, England, and the laboratories of Le Matériel Téléphonique in Paris cooperated in the project. The equipment was largely developed in Paris.

This startling demonstration of the effectiveness of centimeter waves for practical communication over substantial distances and the installations that soon followed it played a significant role in the development of this region of the electromagnetic spectrum.

ITT Federal Laboratories
United States of America

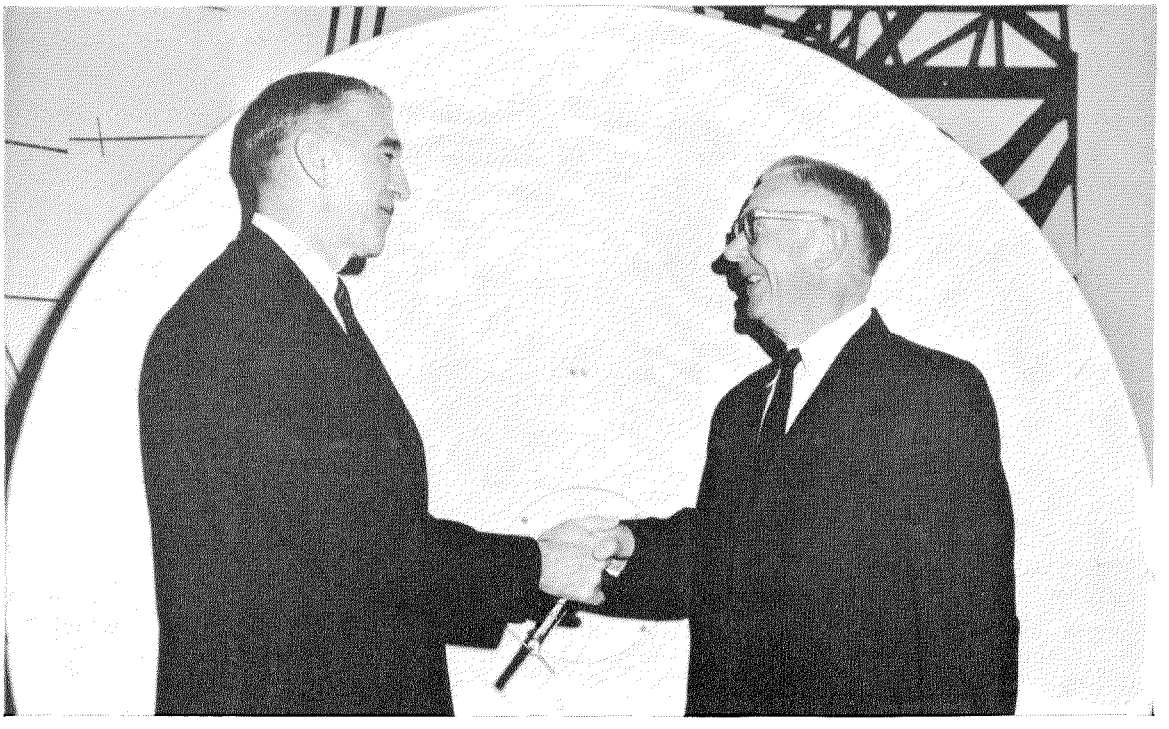


Figure 1—Presentation to Mr. Frank Taylor (at left), director of the United States National Museum, of historical microwave equipment used in the 1931 demonstration of 18-centimeter transmission across the English Channel. The presentation was made by Mr. Andre G. Clavier (at right), who led the engineering team responsible for this pioneering work.

Electronic Nurse—An "Intensive Treatment Telemonitor" was demonstrated at the Third Annual San Diego Symposium for Biomedical Engineering. It is capable of monitoring as many as 25 patients simultaneously for such physiological factors as heart and respiration rates, temperature, and blood pressure.

These factors are presented continually at the bedside and at a central location. They are also recorded permanently. The physician defines the upper and lower limits of each factor as determined by the condition of the patient. The observer will be immediately alerted by both visual and aural means to any significant changes in these factors.

As shown in Figure 2, the system elements at the bedside are the sensors attached to the patient and the electronic data-processing equipment. A transmission system carries the data to a central point, where a display unit keeps the observer informed of the condition of all the patients under supervision, and where a recorder makes a permanent record for each patient.

This system resulted from the development of a monitoring arrangement for continual checking of the physiological conditions of pilots and astronauts.

ITT Federal Laboratories
United States of America

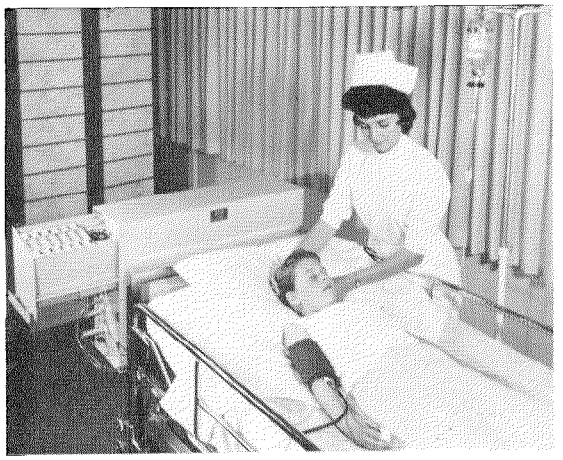


Figure 2—Patient "wired" for the Intensive Treatment Telemonitor. The data processor is in the cabinet at the head of the bed. Heart and respiration rates, temperature, and blood pressure are monitored continually.

Impulse Generator with Program Control—The impulse generator shown in Figure 3 will produce two independent series of impulses. In each series, both the pulse and space durations may be independently adjusted between 50 microseconds and 10 seconds.

Program control permits each series to be produced continually or to be limited to any set number of pulses up to 25. Up to 5 such groups of pulses, with desired spacing between the groups, may be repeated. Also, the two sources, using different patterns of pulses and spaces, may transmit such groups alternately.

Two output circuits, an electromechanical relay and a transistor, are each capable of handling 500 milliamperes at 48 volts. The sensitive polarized relay produces less than 2 milliseconds of distortion and is suitable for relatively low-speed switching, operating effectively to above 100 bauds.

The transistor output operates over the entire working range. Using it, circuit interruptions, vibration, and other disturbances can be simulated. Logic, memory, supervisory, automation,

Figure 3—Impulse generator incorporating program control.

and remote-control systems may be tested with the generator.

Standard Téléphone et Radio
Switzerland

Spark Erosion Machine—The removal of metal by the eroding effects of an electric spark permits the shaping of pieces to highly precise dimensions. The machine shown in Figure 4 was developed originally for laboratory research but its effectiveness prompted its immediate use in manufacturing microwave components and in preparing masks for critical thin-film evaporation techniques.

The machine can cut slots as narrow as 10 microns in material 50 microns thick and slots of 50 microns in 1-millimeter-thick material. The accuracy and resolution is around ±2 microns.

Standard Telecommunication Laboratories
United Kingdom

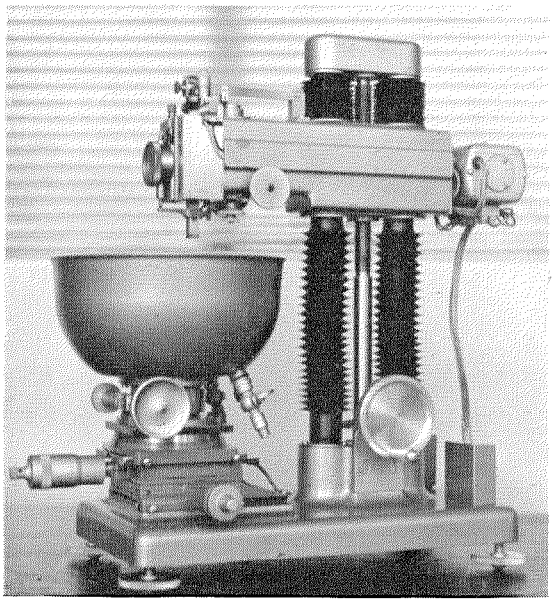


Figure 4—Experimental machine for precise shaping of metal by spark erosion.

Bank Signatures Transmitted by Videx—In 20 seconds, a slow-scan television system permits the central signature file of the Provident Tradesmens Bank and Trust Company in Philadelphia, Pennsylvania, to transmit over a regular telephone line a picture of the signature of a depositor who wishes to transact business at a distant office of the bank.

On receipt of a telephoned request to the central file, the card bearing the signature is placed in a Videx transmitting cabinet and on pressing a button is transmitted over the telephone line to the inquiring office. There, as shown in Figure 5, it appears as a still picture on the screen of a display tube, where it can be retained for several minutes. Operation of an erase button clears the tube immediately.

It is impracticable to ask every depositor in a large banking system having dozens of offices to sign a card for each office or to maintain in each office a file of signatures of depositors in all offices. Videx provides a simple, direct, and rapid means of verifying signatures so that every depositor may be protected in transacting his business at any office of the bank.

ITT Industrial Laboratories Division
United States of America

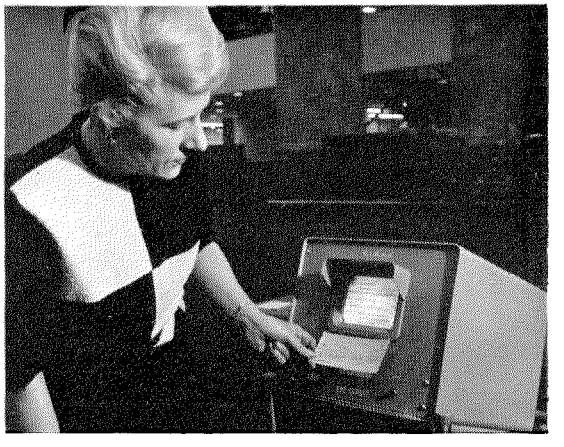


Figure 5—Bank teller verifying signature against original signature card reproduced over telephone line from central file using Videx system.

Reflex Klystrons for Millimeter Waves—A series of reflex klystrons has been developed to cover wavelengths from 3 to 12 millimeters for use as local oscillators and as pump oscillators for parametric amplifiers and masers. Figure 6 is a cutaway view showing the structure.

The 8-millimeter tube has produced 2.3 watts of output power although the production units will be rated between 0.5 and 1 watt. This output is substantially larger than is available from present commercial klystrons in this wavelength band. The sturdy structure results in good frequency stability. Life tests have now been run for only 2000 hours; a much-longer life is anticipated. The assembly is self-jigging, which reduces manufacturing costs.

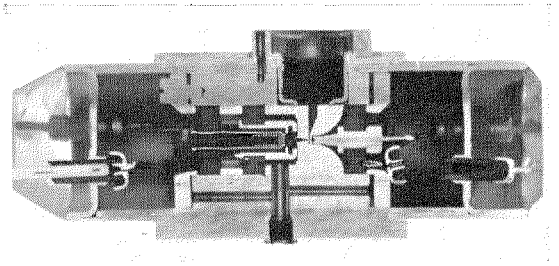


Figure 6—Reflex klystron for millimeter wavelengths. This tube, which operated near 3 millimeters, is about 4 inches (10 centimeters) long.

This klystron is tuned by an external cavity, the resonant frequency of which is controlled by an adjustable multilayer slab of dielectric material. Tuning ranges greater than 3.5 percent of the operating wavelength have been obtained.

Standard Telecommunication Laboratories
United Kingdom

Cheekphone for Noisy Places—To avoid the normal pickup of noise by the microphone of a telephone handset, a short-handled set has been devised that operates with the microphone pressed against the speaker's cheek. A 70-decibel transistor amplifier mounted inside the set amplifies the microphone output.

As shown in Figure 7, the shorter handset hangs on the gravity switch at the back of the pedestal. A shielded cord avoids crosstalk, which could be bothersome with the high amplification that is employed. Power is obtained from the alternating-current mains and a selenium rectifier assures proper polarity of the output voltage.

Fabbrica Apparecchiature per Comunicazioni Elettriche Standard
Italy

Steel-Jacketed Picture Tube—A television picture tube having the critical area of mechanical

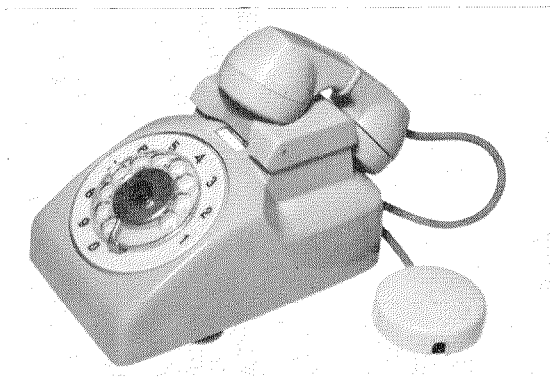


Figure 7—Cheekphone set using short handle to permit microphone to press against the cheek of the speaker and thus exclude ambient noise.

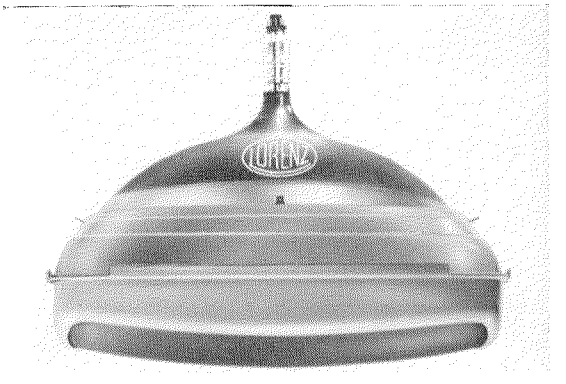


Figure 8—Steel-jacketed television picture tube that needs no safety glass and may project beyond the front of the cabinet.

stress of the glass bulb bonded by a filler material to a steel jacket avoids the need for a transparent safety shield.

The reflecting surfaces of the safety glass are avoided as is the accumulation of dust between it and the face of the tube. The protected outer edge of the bulb may safely extend beyond the cabinet front. The conventional mounting strap is eliminated by flanges on the steel envelope that facilitate both original installation and replacement. Figure 8 shows the *A59-12W* tube with 110-degree deflection and a screen diagonal length of 59 centimeters (23 inches). The over-all length is 36 centimeters ($14\frac{11}{64}$ inches).

Standard Elektrik Lorenz
Germany

Forest Fire Simulator—To train forest service personnel in combating large-scale fires, the simulator illustrated in Figure 9 has been developed. Designed to reproduce the appearance of a forest-fire command post, the system accommodates 12 trainees who face an 8-by-12-foot (2.4-by-3.7-meter) screen on which the fire scene is projected as if it were several miles away.

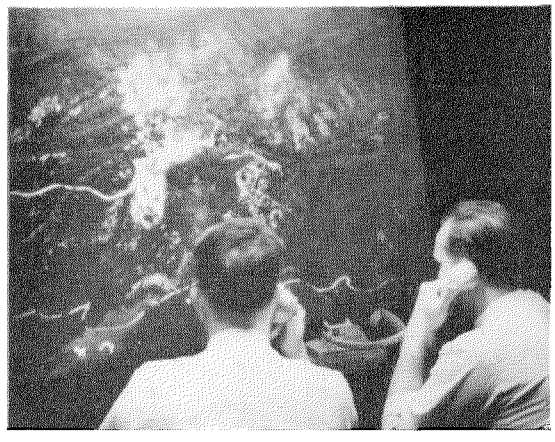


Figure 9—Simulator of a command post with projected image of a distant fire used to train forest service personnel in logistics and techniques of fighting large-scale fires.

The instruction staff consists of an effects operator, a role player, and an umpire, who sits in an isolated booth. The umpire establishes for each exercise the weather conditions, types of fuel in the path of the fire, and other pertinent and usually complicating factors. The effects operator controls the projected view of the fire area. The role player, taking his cues from the umpire, simulates the outside world of the weather bureau, aerial support, police, and other significant agencies.

The trainees must cope with each changing situation by controlling helicopters used for observation, fire-fighting aircraft and trucks, and personnel through orders sent over various communication systems at their disposal. The umpire evaluates the results of all factors and instructs the effects operator to change the projected image accordingly.

Figure 10—Combined phase-inverter triode and push-pull pentodes *ECLL 800*.

An exercise may, of course, include interference with the arrival of needed equipment by heavy road traffic of people interested in seeing the fire, the presence of summer campers in the fire area, and other complications.

International Electric Corporation
United States of America

ECLL 800 Twin-Pentode–Triode—A phase-inverter triode and two push-pull pentodes mounted in a single glass envelope with a 9-pin base simplify the audio system for low-priced radio receivers and phonograph amplifiers.

Class-*B* operation at 250 volts with only 28 milliamperes for the entire tube produces a power output of 6.5 watts with less than 1 percent of harmonic distortion and 9 watts at 5-percent distortion. In *AB* operation, 8.5 watts may be produced with 5-percent distortion and 8 watts with 3-percent distortion. The tube is shown in Figure 10.

Standard Elektrik Lorenz
Germany

7-Gigacycle-Per-Second Radio Link Equipment—Up to 300 voice channels are available over the radio transmitter and receiver, shown in Figure 11, which operates in the 7-gigacycle-per-second region.

Klystrons are used as oscillators in the transmitter and receiver and all other active elements are transistors. Printed boards and plug-in units conserve space and both equipments are mounted on a standard telephone bay 1.8 meters (5.9 feet) high. Working and stand-by sets, automatic change-over equipment, and the service telephone channel can all be accommodated in two bays.

There is provision for connecting the transmitter to the receiver through a frequency converter for adjustment purposes. The antenna filters are placed in the transmission lines and are not mounted in the bays.

Fabbrica Apparecchiature per
Comunicazioni Elettriche Standard
Italy

Color Data Display System—A modified teletypewriter at the Marshall Space Flight Center in Huntsville, Alabama, was used to interrogate an information processor in Paramus, New Jersey, 1000 miles (1600 kilometers) away. In reply, the requested data were sent over an ordinary telephone circuit and projected in seven vivid colors on a screen within half a minute.

The request for information was transmitted to an *ITT 7300 ADX* automatic data transmission system, a real-time stored-program information

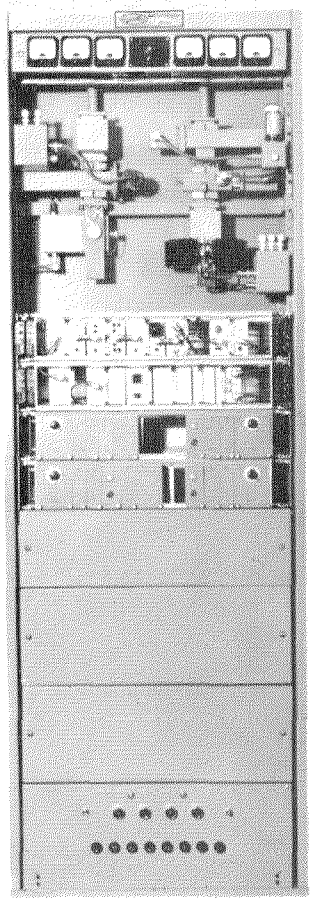


Figure 11—Radio link equipment operating at 7 gigacycles per second and providing for 300 voice channels.

processor suitable for on-line telegraph operation. The reply was transmitted to Huntsville over a Bell Telephone System newly developed *402* data set, which provides for 8 levels of binary information in parallel at 600 bits per second over a conventional voice-grade telephone channel.

The digital data were converted to visual alphanumeric form in seven colors in a Datachrome color projector. The Datachrome projector uses black-and-white 35-millimeter film. The seven colors are obtained by combining the three prime colors obtained by separating a white-light source with mirrors. The information may be projected as far as 50 feet (15 meters).

International Electric Corporation
ITT Federal Laboratories
United States of America

Electromechanical Filters—Electromechanical filters providing Butterworth response in the range from 50 to 500 kilohertz are useful in multiplex telephone and telegraph systems as well as in intermediate-frequency amplifiers and in single-sideband radio equipment.

A typical filter is shown in Figure 12. The design offers high performance in small size. Stability is assured by the choice of materials.

Standard Telecommunication Laboratories
United Kingdom

Recorder for Crossbar-Switch Utilization—Continual observations of the busy condition of crossbar selectors will expose overloads and permit such assignment of lines and switches as provides maximum service to all subscribers. Accurate estimates may also be made of needed expansion of an exchange.

An electronic analyzer permits each element to be checked during a 20-millisecond scan and its condition reported to a ferrite memory. The memories are then read and the number of seized elements in each frame is indicated on a counting relay or reported on a page teleprinter. A simultaneous operation reports on counting relays busy information on each of the 320 elements that may be observed. The observation periods can be controlled either manually or automatically.

The equipment has been used in Pentaconta public telephone exchanges to study traffic distribution in secondary selectors (links) and final selectors at the subscribers stage.

Compagnie Générale de Constructions Téléphoniques
France

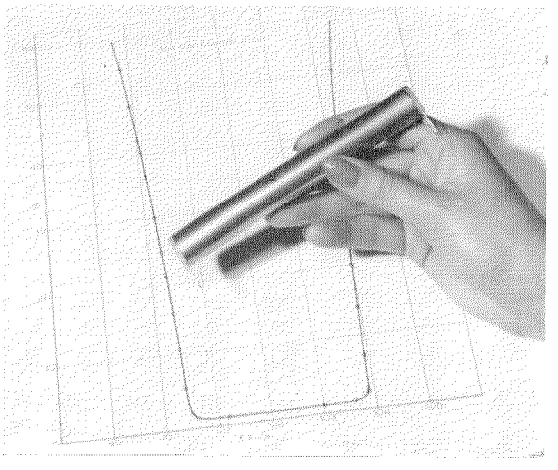


Figure 12—Electromechanical filter producing Butterworth response.

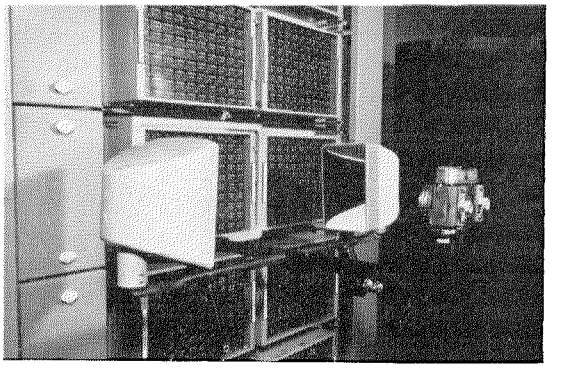


Figure 13—Subscriber meter frame with photographic equipment attached for recording calls for billing.

Telephone Subscriber Meter Frame—A new meter frame has been designed for Pentaconta exchanges. Two groups of 100 subscriber meters each are mounted on the frame and 10 such frames, containing 2000 meters, are accommodated in a standard Pentaconta bay. An identity finder frame is located in the lowest part of the bay.

Figure 13 shows part of a bay of meters to which photographic equipment has been attached for recording the readings for billing purposes.

Manufacture has begun for installation in the Pentaconta exchanges of the Compañia Telefonica Nacional de España in Spain.

Standard Eléctrica
Spain

Polyethylene-Terephthalate Dielectric Capacitors—Films of polyethylene-terephthalate (Mylar) as thin as 3.8 microns are available for use as dielectric in capacitors. Even a single layer may be used if the metal electrodes are applied by evaporation. They then exhibit the self-healing properties of other metallized paper capacitors. These polyester materials absorb very little moisture and capacitors need not be expensively sealed to ensure adequate life.

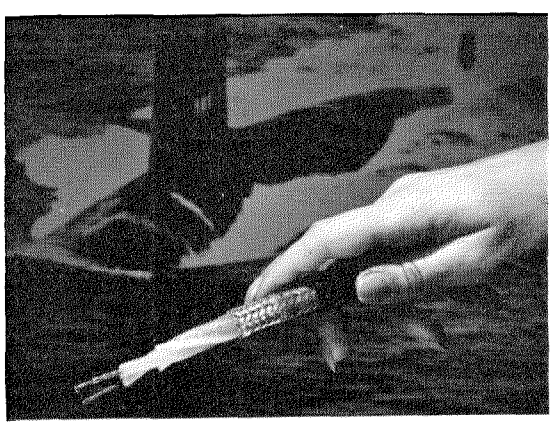


Figure 14—Multiconductor cable for use in submarines has all interstitial voids filled with a compound to prevent leakage through the length of a cable that may pass through bulkheads in the ship.

Series of capacitors are available in both simple tubular form and in metal cases. Capacitance ranges from 0.1 to 47 microfarads at direct-voltage ratings between 100 and 630 volts. The operating temperature range is from -55 to $+125$ degrees centigrade. At room temperature, the insulation resistance is between 5000 and 10 000 ohm-farads and at 1 kilocycle per second the loss angle, tan δ, is 0.8×10^{-3}.

Standard Elektrik Lorenz
Germany

Line Concentrator and Telephone Carrier—A *K-52* line concentrator that would normally serve 52 subscriber lines over 14 physical pairs of wire has been combined with a *K-31* rural carrier system to further reduce the required wire lines to 1 or 2 pairs.

Line concentrators provide simultaneous service to a number of randomly selected subscribers' lines that terminate in a relatively small geographical area over a much-smaller number of lines to the central office. This is feasible because residential subscriber lines average less than 10-minutes use per day.

The carrier system removes any limitation on the loop resistance between the concentrator and the central office. The limit for the *K-52* concentrator would be about 1400 ohms or 7 miles of a 22-gauge pair in a cable.

The new line concentrator and carrier has been successfully field tested by the Allied Telephone Company in Sheridan, Arkansas.

ITT Kellogg Telecommunications
United States of America

Radio Altimeter—The British Overseas Aircraft Corporation is installing the *STR.51* altimeter in its fleet of *VC-10* airliners. An essential element in all-weather landing operations, its accuracy at low altitudes is such as to make it useful over both the flare-out and touch-down phases of landing.

Duplicate units and automatic monitoring ensure prompt changeover in case of equipment

failure. The apparatus may be as much as 50 feet (15.3 meters) from its antenna system.

Standard Telephones and Cables
United Kingdom

Water-Blocked Multiconductor Cable—A multiconductor cable has been developed for use in submarines that, if cut, will not leak water through the interstices running parallel to the conductors in the cable. An impermeable viscous compound fills all these interstices without affecting either flexibility or electric properties. The cable will not leak even when subjected to very high hydraulic pressure.

A typical design is shown in Figure 14. The cables are available in coaxial form with copper braided shield and plastic jacket or in shielded individuals, pairs, and triads of up to 132 conductors.

Surprenant Manufacturing Company
United States of America

North Sea Telephone Cables—The British Post Office in conjunction with the German, Danish, and Netherlands administrations has ordered the 4 new submarine-cable systems shown in Figure 15. All are capable of providing 120 telephone channels, and the Germany and Denmark links will be the first direct high-capacity telephone cables to connect these countries with England. Two separate cables will be laid to Germany.

Over 900 nautical miles (1667 kilometers) of conventional armored cable will be required. Most of 0.62-inch (15.7-millimeter) diameter, a section between Borkum Island and the German mainland will be of 0.935-inch (23.7-millimeter) size to avoid the need for repeaters in very shallow water.

These will be the first installations of new low-cost shallow-water repeaters of which there will be 81 and 6 equalizers. The demountable housings for these units have been described [1].

Standard Telephones and Cables
United Kingdom

1. B. M. Dawidziuk and F. L. Jarvis, "Shipboard-Adjustable Submerged Equaliser," *Electrical Communication*, volume 38, number 1, pages 88–105; 1963.

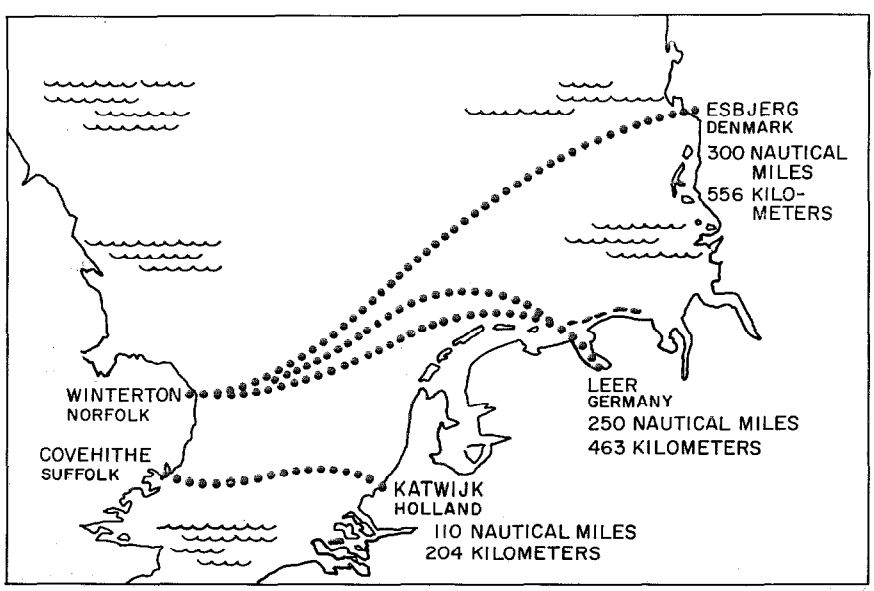


Figure 15—Map of 4 new submarine telephone cables to be laid across the North Sea.

Assistant Telephone Set for Call Assignment and Handsfree Operation—An executive's telephone station permitting handsfree operation is shown in Figure 16. It is based on the Assistant type of subscribers equipment. Provision is made for call assignment, which permits transfer of a call to the executive's set from that of the secretary.

Either handset or handsfree operation may be used, the latter employing a transistor amplifier and a suppressor to eliminate the damaging effects of electroacoustic feedback. The amplifier operates from the direct-current supply from the central office.

With the exception of the loudspeaker, which is in a separate housing, all components are mounted in the subscribers set and in an additional small box. A second microphone, also for handsfree operation, may be connected to the desk set.

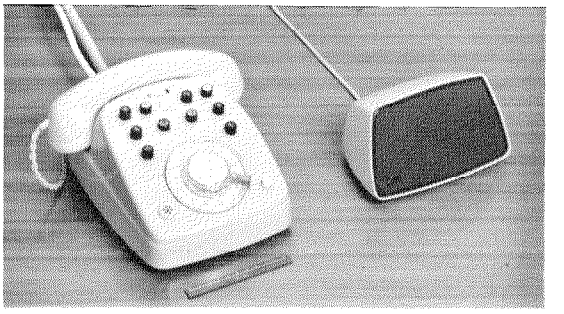


Figure 16—Assistant type 2-1 executive's subscriber set and separate loudspeaker for handsfree operation.

Standard Elektrik Lorenz
Germany

Trends in Component Developments

P. H. SPAGNOLETTI
Standard Telephones and Cables Limited; London, England

1. Introduction

In the advanced countries of the world, electronics is one of the fastest growing industries. A main spur behind this growth has been the continuous development of new products, many of which are available only because of progress in components development. Radar would not have been possible without high-power vacuum tubes operating in the centimetre band. Modern computers, with their high speed and capacity, are practical because of the low power consumption and small physical size of semiconductors. Microwave links owe a lot to the travelling-wave tube, and the extremely high reliability of special components has made feasible the use of submerged repeaters.

This short survey of electronic components is presented from a European point of view, although it is inevitably coloured by the experience and philosophy of the company with which the author is associated.

Since its early days, after the First World War, when it was concerned largely with radio broadcasting, the electronics industry has grown enormously; in the United States some 35 times since 1939. To-day, it is one of the largest in the country. Similar rates of growth have been experienced in the United Kingdom and other European countries. There have been three stimuli during the growth period. To some extent these have overlapped in time. Firstly, radio broadcasting and the following television development; secondly, military electronics and radar; and finally, professional electronics including communication.

The radio and entertainment industry has already lost its dominant role in the United States, where it now represents only a quarter of the total sales in electronics, and has similarly fallen in the United Kingdom to about one third of the total turnover. In both these markets, professional and defence products have now a greater importance.

The three phases or stimuli in the history of electronic components have had quite a marked influence on the products themselves. In the entertainment industry, price has always been the dominant feature, and, as a result of price competition, has come mechanization. The military market requiring, as it did during the Second World War, large quantities of components looked mainly to the entertainment-equipment factories for this production. However, the military superimposed its additional requirements for severe environmental conditions and for a lower failure rate than was acceptable in the entertainment market. Close tolerances were not essential nor was long life; the entertainment specification for the latter was traditionally 2000 hours, whilst the working life of components in missiles, including ground testing, could be as low as 200 hours; shelf life, of course, here had to be excellent.

At the same time as the military market was going through this most rapid phase of expansion, the telecommunications and electronics markets were requesting specialized components having failure rates as low as, or lower than, those in the military field, but with close tolerances to reduce the engineering costs of systems made in comparatively small quantities. The primary requirement of these components was, however, long life, of the order of 5 to 10 years in most systems; while, for submerged repeaters, lives had to exceed 20 years of continuous operation; environmental conditions, on the other hand, were less stringent than for military requirements. To-day, the tendency is for military and electronics requirements to coalesce. With the increased need for ground control equipment and the wider applications of computers, the longer-term aspects of reliability are increasing in importance. The techniques of submerged-repeater manufacture, where each component has its own pedigree, a punched-card record of its various stages of manufacture, and assembly is undertaken in scrupulously clean conditions, have been used for the components for the Minuteman project.

2. Growth

The high rate of growth of the electronic industry, as was mentioned earlier, would not have been possible without the introduction of new products. New products that not only fulfilled needs that were previously satisfied by other means, but also needs that had been awaiting technological advances before their satisfaction was a practicable proposition. For example, analogue computers for forecasting upper-atmosphere developments have made possible the complicated calculations needed to improve weather forecasting. The time element was sine qua non. Banks are now using electronic data processing for maintaining customer accounts and cheque sorting in place of clerical effort, whilst many production control processes are now handled by electronic equipment rather than by operators. In the military field, the importance of these new products may be shown by the increasing share of the defence dollar absorbed by electronics. In the United States, electronics accounted for 6·5 per cent of military procurements in 1951 but 18 per cent in 1961. That this growth is expected to continue is shown by estimates that expenditure on research and development in this field, which represented 35 per cent in 1959, will grow to 50 per cent in 1965 and 65 per cent in 1970.

For many years the growth of the components sector was in line with that of the equipments using these components. Recently, however, the components growth has exceeded equipment growth for two reasons. Firstly, as components become smaller, the supporting structures diminish in size and cost, and, with the use of the printed circuit board and multiple soldering with automatic assembly, the labour costs of equipments are reduced; components themselves, therefore, represent a larger part of the total cost. Secondly, these electronic components are also now being widely used outside the industry. Transistors and associated components are being used in watches, toys, and in automobiles for ignition and power systems. The development of semiconductors, the decrease in size of the active components, and the corresponding decreases in size brought about by the lower voltage and current requirements of the passive components has enabled highly complex equipments to be practicable. The manufacture and assembly of the vast numbers of components for such equipments has been aptly called the "tyranny of numbers" by Jack Morton of Bell Telephone Laboratories. As an example, the number of electronic components in military aircraft has increased 10 times between 1950 and 1960. Such large numbers bring home very forcibly the reliability problem, as a simple argument will show.

3. Reliability [1]

If one component in 10 000 fails within a given period, say 1000 hours, that is, a failure rate of 0·01 per cent or a reliability of 0·9999 per 1000 hours, an equipment of 100 components would have a reliability of 99 per cent; in an equipment of 1000 components, the reliability falls to 90 per cent; and with 10 000 components it is only 37 per cent. Thus, after 1000 hours, on only 37 per cent of the occasions when service is required would the equipment operate satisfactorily. Modern telecommunication equipment requires that the component failure rate shall not exceed 0·01 per cent per annum. In submerged repeaters, where the recovery and replacement cost of a component is prohibitively high, A. A. New [2] of the British Post Office quotes failure rates of 0·0003 per cent per annum as essential. Figures approaching this are also becoming vital for very-large installations.

Such reliability must inevitably raise the cost of these components, but they are still economic when weighed against the alternatives of redundancy or high maintenance costs. Even where the equipment may be said to be readily available, such as in aircraft, the Royal Air Force estimate maintenance costs to be 3 to 4 times the cost of the original

equipment. Component reliability has unquestionably improved over the years but is only just keeping pace with the needs. Fifteen years ago, reliable components were simply selections from the normal commercial production, principally for the entertainment industry. But it was very soon realized that you cannot "inspect quality into a product."

To-day, such components are specifically designed and manufactured in closely controlled conditions with inspection at every stage. Figure 1 shows a typical "extra-clean area." Mechanical or automatic processes, which are so essential for consistently higher reliability, have reduced the effects of human limitations. But where operators are necessary, experience has shown that they must be willing to make a positive contribution to reliability in manufacture; and, for this purpose, integrity is all-important: admitting carelessness or acknowledging limitations of knowledge will often prevent a faulty component from being made. As a measure of the improvements that have been achieved, the following typical figures are quoted from Mr. New's publication [2]. In British submerged repeaters, the failure rate per annum in 1951 was 0·08 per cent, whilst in the period 1954–1959 it was an order of magnitude better at 0·004 per cent. Valves alone had improved from 4·1 per cent to 0·20 per cent.

These low level rates create, obviously, a problem of measurement, which is examined in detail in another article [1] published in this issue. In this latter article, the example relating to the application of the Arrhenius equation to the ageing law of germanium transistors as a function of temperature is particularly illustrative. It results from an investigation published by J. M. Groocock and R. H. Chilton [3]. It may be seen that extrapolated times-to-failure for a temperature of 45 degrees centigrade have to be expressed in hundreds of thousands of hours. These results, incidentally, demonstrate the improvement obtainable by operating at reduced temperature.

Other life acceleration tests, using the step-stress technique, can produce similar life data under various electrical and mechanical stresses and provide as complete a picture as possible of the failure mechanism of a device. As a further example, the paper capacitors used in submerged repeaters are subjected to a stress of 6·5 kilovolts and an ionization discharge count is made to indicate the presence or otherwise of small voids in the impregnant: voids that contribute to a shortened life. For this particular application, it is essential that there should be no voids at all.

4. Next Phase in Circuits

The tyranny of numbers and the drive for reliability is responsible for the next phase in component developments. Interconnections or joints between components are fundamental causes of trouble: taking the earlier example again, if it is assumed that each component has two soldered joints, and if one joint in every thousand fails each 1000 hours (that is, a reliability of 0·1 per cent per 1000 hours—a figure often quoted) the reliability of the equipment with 100 components will drop to 81 per cent and that of an equipment using 1000 components to only 12 per cent. All connections or interfaces of different materials are potential trouble spots, and even the simplest components, such as cracked-carbon resistors, will have as many as 8 such connections or interfaces when they are mounted in a circuit.

It is obvious, therefore, how essential it is to reduce such interfaces to the absolute minimum. Two lines of attack are under intensive development throughout the world. The semiconductor circuit, where a single chip of silicon or other semiconductor material has built into it the active components required, as well as the resistors and capacitors; and the thin-film circuit, where a substrate of glass has deposited on it capacitors and resistors to which the active elements are assembled subsequently. Both of these approaches enor-

Figure 1—Typical extra-clean area for assembly of components.

mously reduce the number of interfaces and connections, and thus will contribute materially to reliability. This has, however, yet to be proved statistically. Apart, however, from the reliability advantages that such thin-film and semiconductor circuits give us, there is the additional requirement—for nanosecond computers; a signal will travel only 10 inches in this unit of time, so that the distance between the component parts becomes significant in equipment operation.

5. Semiconductor Circuits

Most of the semiconductor circuits available to-day perform specific digital functions, for it is in digital applications that one is likely to require large quantities of similar circuits. Only by the production of large quantities can the price fall to the economic levels that make these devices competitive with existing methods. However, a number of authors have questioned whether this is the right approach. Rice [4] mentions an attempt to build a

computer using 8 standard blocks; he reported that such a computer was feasible but that it would have better performance at lower cost if 486 circuits were used. Inevitably, standard blocks must involve compromises, in which case they are unlikely to be the optimum design for any particular application—so performance must be sacrificed. The performance of a system depends not only on the effective operation of the individual components but on the clever circuit design that makes optimum use of their characteristics.

The industry has been trying with very limited success to standardize its components for many years. It seems hard to believe, therefore, that it would be willing to accept standard circuits. J. M. Golday of Bell Telephone Laboratories suggests that better system performance, flexibility, and lower costs are achievable if only similar devices are fabricated on a common substrate; a number of substrates being put in a common encapsulation to form a circuit if desired. Golday maintains that "the major advantage of semiconductor circuits comes from the use of a common package, not a common substrate." Thin-film circuits, however, can be made economically in quantities and thus may overcome (at least for a period) the disadvantages imposed by the limitations of standard building blocks and the dislike of the average designer to be tied to standards, no matter how good they may be.

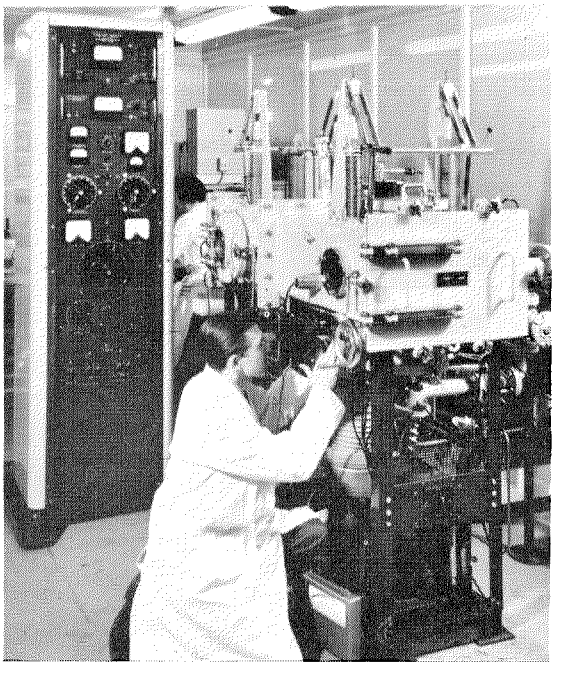


Figure 2—Equipment for producing complete thin-film circuits of passive components.

6. Manufacturing Techniques

The adoption of either semiconductor circuits or thin-film circuits raises two immediate problems. For component manufacturers, it breaks down the present divisions of products. For the equipment manufacturer, it means either increasing dependence on the suppliers of these two products with the concomitant loss of a further portion of a share of the value added in manufacture or a necessity to integrate vertically by setting up his own thin-film or solid-state factory. The expense and know-how required must, however, make this a risky and possibly undesirable process for many equipment manufacturers. It is obvious that the close liaison that exists to-day between the equipment and component designer will be even more essential in the future.

The new components, in whatever form, are inducing quite fundamental changes in the requirements for men and facilities in the production processes. Until recently component manufacture was basically an assembly operation. Close control of the properties of the parts used, such as metal parts in valves and paper and mica sheets in capacitors, was not essential to enable the product to function satisfactorily. Empirical knowledge and experience were often sufficient. To-day, however, the physicist and the chemist are far more important to the components industry. Semiconductors require materials of exceptional purity and close control or modification of their crystal structure for their satisfactory operation. The surface structure of substrates

for thin films and the effect of its life on the resulting assemblies is of great importance.

Most components can be made by thin-film techniques, and others are under development. Figure 2 shows equipment capable of producing a complete thin-film circuit of passive components without the necessity for removing the substrate. Metal and metal-oxide resistors are used both as individual components or as parts of circuits. Thin-film memory planes are challenging the toroidal memory devices.

Such changes and the overriding need for reliability are putting new requirements on the production floor. Mechanization, which used to be thought of mainly as a cost-reduction process, is to-day required to attain uniformity and reliability. H. H. Arnold [5] of the Western Electric Company says, "automatic control over manufacture can make it possible to produce a product at a higher level of reliability than would be attainable otherwise." He demonstrates the substantial improvements (of at least four

Figure 3—Cabinets providing small, clean, temperature-and-humidity-controlled enclosures in which transistors are encapsulated.

times) in the reliability of carbon-film resistors which he attributes to the use of computer-controlled automatic processing.

Again for reasons of reliability and also because of size reduction, modern component factories must be laid out to facilitate incorporation of dust-free temperature-controlled and humidity-controlled areas. The size of dust particles is now significant when compared with the product; measurements are now recorded in microns and Ångströms. It is not necessary, nor is it convenient, for the whole factory to be so treated, but small areas or even special cabinets must be readily available. Figure 3 shows a typical cabinet used to encapsulate transistors.

The rate of development in transistor fabrication techniques has given rise to its own manufacturing problems, in that a technique has often been superseded by the time production was in full swing. There seems to be general agreement, however, that some stabilization is likely with the planar epitaxial approach. The main drive leading to this development has, of course, been the desire for transistors of higher and higher performance, but in parallel has been the search for devices that lend themselves to lower-cost production and which are inherently reliable. Efforts to improve the reliability of germanium alloyed transistors have met with considerable success with the introduction of the molecular sieve and the cold-welded case. Figure 4 shows the improvement in life obtainable over earlier methods of encapsulation.

7. Future Developments

To-day's communication systems operate basically between direct current and 10 gigacycles per second. The extension of a further octave, therefore, gives us 9 times the present available bandwidth. It is in this octave, between 10 and 100 gigacycles per second, that the vacuum-tube engineers are working. Developments made possible by accurate machining techniques using, say, spark erosion or an electron beam, have extended the use of microwave tubes to these frequencies, and even beyond. Klystrons, backward-wave oscillators, and crossed-field devices have been developed, giving tens of milliwatts at these frequencies. Indications are that it may be possible to develop travelling-wave tubes capable of 10 to 15 watts of continuous-wave power at frequencies between 50 and 100 gigacycles per second.

Following the development of the parametric amplifier with its almost theoretical noise factor, making satellite communication a practical proposition, interest has jumped to the infra-red and light regions of the spectrum (300 000 gigacycles per second). Coherent light in the form of the laser is now practicable and ways of overcoming the problems of continuous-wave operation and modulation

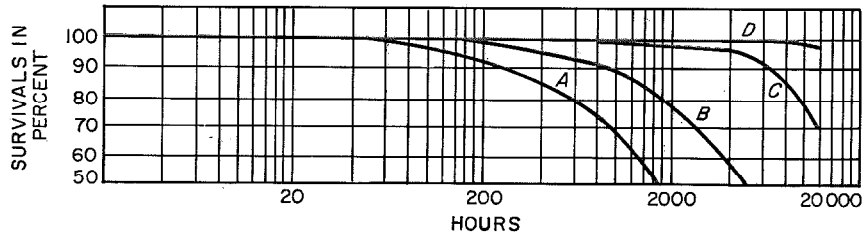


Figure 4—Survival curves for germanium alloy transistors. Curve *A* is for temperature-soldered cases and the semiconductor surface covered with silicon grease stored at 65 degrees centigrade. Curve *B* is in an araldite moulded case, temperature soldered, and stored at 65 degrees centigrade. Curve *C* is for a cold-welded case without an absorber and stored at 75 degrees centigrade. Curve *D* is for a cold-welded case, a molecular sieve built-in adsorber, and storage at 75 degrees centigrade.

are under development. One of the most promising approaches appears to be a further development of the gallium-arsenide light-emitting diode. This should enable very-much-higher efficiencies to be achieved in coherent light compared with the 1 or 2 per cent for ruby crystal or gas lasers.

Vacuum deposition will provide many new devices for the future. As mentioned previously, thin-film memories, thin-film circuits, and thin-film passive components will find increasing applications to-morrow. Because of the belief that such thin-film circuits, on account of their flexibility, may hold one of the best promises for the future, the development of the corresponding thin-film active devices is receiving a great amount of attention.

Another line, which is already producing some important devices, is the use of molecular excitation for controlling frequencies and for measuring magnetic effects. These have made possible accurate measurements of time, which could lead to new or improved navigation or guidance systems. Cathode-ray tubes with spot sizes of one third of a mil and capable of a resolution of 3000 lines per inch are being used for information-retrieval systems. The whole range of possible devices that are based on the low resistances of metals and alloys at temperatures near absolute zero are still only in the embryonic stage.

Further efforts are continuously being directed to extending the maximum frequency of operation of semiconductors and their power-handling capabilities. The upper frequency limit achievable to-day is of the same order as that reached by vacuum tubes using the conventional cathode, grid, and anode structures, although the outputs are somewhat lower. To overcome inherent limitations, vacuum-tube engineers had to employ new principles, the slow-wave structure and the cavity. It is likely that the semiconductors that will eventually compete with these vacuum-tube devices will also depart from the accepted emitter, base, and collector structure. Pierre Aigrain [6] has described active homogeneous semiconductor devices; that is, devices without junctions that do not depend on minority carriers for operation. These show great promise at high frequencies and high power. The encapsulation of components has advanced from the wax and paper era when radio and television provided the main market, through the glass, hermetic seal, and epoxy era of military components, and now to the planar transistor or diode that in many applications may not require encapsulation at all to protect it from its environment. Figure 11 in the article by B. D. Mills [7] demonstrates the stability of h_{fe} in a sample of 24 unencapsulated silicon planar transistors during a life test that has already lasted over 7000 hours. Other component developments, such as the silicon-dioxide capacitor and the metal-film resistor, place a lower requirement on the encapsulation, and are, therefore, potentially more reliable.

Various units of the International System in Europe and America have been and are in the forefront of many of these components developments. The impetus behind these has generally stemmed from the need and thus the desire to produce components with greater reliabilities, longer life expectations, and novel properties for the progress of the electrical communication art.

8. References

1. C. A. Meuleau, "High Reliability Testing and Assurance for Electronic Components," *Electrical Communication*, volume 38, number 3, pages 307–324; 1963.

2. A. A. New, "Reliable Components in Post Office Equipment." Presented at the Second Symposium on Electronic Equipment Reliability at the Institution of Electrical Engineers in London on 24–26 October 1962.

3. J. M. Groocock and R. H. Chilton, "The Use of Acceleration Factors and Failure Distributions in the Prediction of Transistor

Failure Rates." Presented at the Second Symposium on Electronic Equipment Reliability at the Institution of Electrical Engineers in London on 24–26 October 1962.

4. Rex Rice "A System Designer Views Micro Integrated Electronics." Presented at the Solid State Device Research Conference in Palo Alto, California, on 26 June 1961.

5. H. H. Arnold, "Reliability: A Benefit of Automation," *Western Electric Engineer*, volume 6, pages 4–10; January 1962.

6. P. Aigrain, at the 1962 International Solid State Circuits Conference held in Philadelphia, Pennsylvania.

7. B. D. Mills, "Silicon Epitaxial Planar Transistors—Part 1—Appraisal," *Electrical Communication*, volume 38, number 3, pages 363–371; 1963.

P. H. Spagnoletti was born on 6 June 1906 in Sydenham, Kent, England. He was educated at Lancing College and at Trinity College, Cambridge, where he obtained a National Science Tripos in 1928.

He started his engineering career on a high-frequency world-wide radiotelephone system, working chiefly in England, Spain, and China. In 1931, he was responsible for the planning, engineering, and installation of the first two empire broadcasting stations for the British Broadcasting Corporation.

He was manager of the airborne radio communication equipment work of Kolster-Brandes throughout the war. In 1945, he became chief engineer and in 1947, general manager of the company, including manager of the valve activity. In 1958, he was appointed components group manager being responsible for all components within Standard Telephones and Cables. He serves also as a director of Standard Telecommunication Laboratories.

Mr. Spagnoletti is vice-president of the Television Society. From 1951 to 1954, he was chairman of the British Radio and Electronics Manufacturers Association and in this capacity represented the industry on technical and other discussions in connection with television standards and the formation of the new television authority.

Mr. Spagnoletti is an Officer of the Order of the British Empire. He is a Member of the Institution of Electrical Engineers.

High-Reliability Testing and Assurance for Electronic Components

CHARLES A. MEULEAU
International Telephone and Telegraph Europe

The necessarily restricted length of this article does not permit all definitions of terms used in reliability to be reviewed, and therefore only terms that are in current usage will be employed. The reader may refer to certain specialized publications [1]. The symbols used are defined in Section 8.

1. Failure Rates

It is usual to characterize the reliability of a component by a failure rate λ. If, among a very-large (infinite) population of identical components, N of them are still operative after a time t, by definition

$$\lambda = -\frac{1}{N}\frac{dN}{dt}. \qquad (1)$$

A priori, λ is a function of time. It is frequently expressed in the form x failures in percent per 1000 hours or also as $x \times 10^{-5}$ failure per component-hour or, abbreviated, $x \times 10^{-5}$. Thus, 0.1 percent per 1000 hours is equivalent to 10^{-6}.

This definition is interesting because in many practical cases and within certain limits the failure rate is a constant with respect to time. It is then possible to integrate (1) in the form

$$R = N/N_0 = \exp[-\lambda t] \qquad (2A)$$

where N_0 is the initial number of components. R is called the proportion of survivals at time t. If

$$S = 1 - R \qquad (2B)$$

S is the proportion of failures at time t. Again when λ is a constant, it may be demonstrated that the statistical mean life of a component is

$$\bar{t} = 1/\lambda. \qquad (3)$$

This is statistically the time at which 37 percent of the components are still operative.

Effectively, during the life of a group of components, there is a more-or-less-extended phase between the initial "infant mortality" and the wear-out* period where λ is practically constant. In principle, this is the phase referred to when nothing more explicit is indicated. But there is, here, a very regrettable risk of confusion in the suggestion that the rates thus defined could be utilized to design a large equipment in which components operate for such a long time that they are generally in the wear-out phase. Mortality is then much higher. Nonetheless, let us for the moment confine ourselves to the limits of the accepted usage of the term.

2. Direct Measurement of Failure Rates

2.1 Statistical Aspects, Sampling Method

When λ is a constant, it is said that we are dealing with a Poisson process. The following problem then arises: to determine how many samples n should undergo test during a period of h hours so that the observation of m failures assures a failure rate of λ_1 (expected) with a specified probability P also termed the confidence level. The *MIL-STD-105* sampling norm, well known, is inadequate in this case because it is based on attribute sampling (the objects being classified into good and bad). Here the time factor is involved. The mathematical base of a sampling plan in the case of a Poisson law is the following relation

$$P = 1 - \exp\left[-m\frac{\lambda_0}{\lambda_1}\right] \times \left\{1 + m\frac{\lambda_1}{\lambda_0} + \frac{\left(m\frac{\lambda_1}{\lambda_0}\right)^2}{2!} + \cdots + \frac{\left(m\frac{\lambda_1}{\lambda_0}\right)^m}{m!}\right\}. \qquad (4)$$

λ_0 is the observed failure rate for a sampling; in other words

$$\lambda_0 = m/nh \qquad (5)$$

and nh is called the number of component-hours of test. P then is the probability that the

* Wear-out as used here is not limited to mechanical deterioration but includes aging, degradation, and similar weakening of desired properties.

actual rate for the entire lot (supposed infinite) does not exceed the value λ_1 that has been specified.

From this, it is more convenient to deduce that, to assure a rate λ_1 with a confidence level P, the minimum value of nh must satisfy the equation

$$nh \cdot \lambda_1 = k \qquad (6)$$

where k is now a numerical coefficient that depends on P and on m. Its value is given in Table 1.

As an example, suppose we verify a failure rate of the order of 0.1 percent in 1000 hours, or more briefly 10^{-6}, with a 90-percent degree of confidence. The minimum value of nh, anticipating that no failure will be observed, corresponds to $k = 2.3$, and thus to 2.3 million component-hours. Since, by hypothesis, λ is a constant, in theory 230 samples can just as well be observed during 10 000 hours, 1000 samples during 2300 hours, or 2300 samples during 1000 hours, et cetera. Practically, this is obviously not irrelevant and we will come back to this question.

The above represents the minimum test. A slightly deeper analysis shows that it has not a high statistical significance. In fact, by means of Table 1, it may be determined that, having observed λ_0, there is 90-percent probability that

$$0.05\ \lambda_0 < \text{true } \lambda < 3\lambda_0. \qquad (7)$$

It thus appears advisable to tighten this interval by observing several failures. For 2 failures, $k = 5.3$, which leads to 5.3 million component-hours. The ratio between the extreme values of the interval then falls to 8, instead of 60.

It is intended here merely to point out certain aspects of these methods and to fix orders of magnitude. Usual sampling plans are even more complicated. Besides the *MIL-S-19500C* specification, which is known to specialists, the work already cited [1] may be consulted.

No serious difficulties, practical or economic, are encountered as yet with the number of samples determined above for an expected λ_1 of 10^{-6}, as far as carrying out the test is concerned. But the determination of failure rates of the order of 10^{-7} (0.01 percent in 1000 hours) requires 10 times the number of component-hours. It can be estimated that this represents the limit order of magnitude for *directly* testable reliability levels, leading to expensive tests, which would have to be justified by the importance of the envisaged application.

2.2 Physical Aspects

One should not overlook the complex nature of reliability in the sense that direct measurement of λ, as described above, assumes the following factors as defined: the component (including the *manufacturer*), the utilization circuit, and the environmental conditions. For the time being, the question of duration of operation is not being considered. It is obviously impossible to vary all these parameters in multiplying the tests. This leads to the

TABLE 1
VALUES OF k IN (6)

Confidence Level P	Number of Failures Observed m										
	0	1	2	3	4	5	10	15	20	30	40
0.95	3	4.7	6.3	7.7	9.1	10.6	17	23	29	41	52
0.90	2.3	3.9	5.3	6.7	8.0	9.3	15.4	21	27	38	49
0.10	0.1	0.52	1.1	1.7	2.4	3.1	7.0	11	15.4	24	33
0.05	0.05	0.35	0.82	1.35	2.0	2.6	6.2	10	14	22.5	31

separate consideration of two fundamental types of problems, which we are now going to examine.

2.3 Determination of λ for a Known Application

Test conditions are then well defined and constitute a single group. This is the problem treated above. There are indeed very few examples that really illustrate such a case. The following is among the most representative.

2.3.1 *Example 1*

This project to determine the long-term reliability of *2N357* and *2N396* transistors was undertaken jointly by General Electric Company and Radio Corporation of America as part of the Ballistic Missile Early Warning System (BMEWS) program [2]. Approximately 10 000 transistors of each type were tested for 10 000 hours, that is, 10^8 component-hours of testing for each type. The test conditions were varied, so that analysis of the results enabled a correlation to be found between the failure rates and the operating conditions. But the whole of the conditions applied was also considered as representative of the expected operating conditions and over-all failure rates were obtained. The example can be considered as an illustration of the above solely from this latter angle. Finally, these over-all failure rates were evaluated at 0.17 percent per 1000 hours (1.7×10^{-6}) for the *2N357*, and 0.024 percent per 1000 hours (0.24×10^{-6}) for the *2N396*, with a confidence level of 80 percent and a precision of 25 percent. The actual rates observed in operation were 0.071 percent and 0.027 percent per 1000 hours, respectively. The 100 000 perforated cards that recorded the measurements made on the *2N396* alone reveal that such a project is not a minor affair.

These tests of large quantities of electronic components should not be confused with similar ones made to preselect components for unusually critical applications, such as, transistors for artificial satellites and tubes for submerged repeaters. In these cases, the purpose of the tests is to *guarantee* the highest possible reliability through the choice of the most-stable components (for example, 5 or 10 percent of the initial lot) rather than to *measure* the reliability. These latter examples, therefore, are outside the scope of the present article.

2.4 A Priori Determination of Failure Rates

If the question, "What is the average failure rate for dry tantalum capacitors?," is posed, it would seem to have very little meaning in view of the above considerations. It is necessary, however, to try to give it some meaning for two useful purposes.

On the one hand, the user could require this information, even though approximate, to know, for example, whether long-life aluminum electrolytic capacitors, in view of their lower price, would also be as satisfactory for a particular application.

On the other hand, the time has come when the manufacturer must supply data on the reliability of his products, and even guarantees.

The current tendency is to present such information in the form of basic failure rates. These are established for somewhat severe operating conditions, for example, maximum temperature and maximum voltage, so that their value, thus increased, can be more easily and more economically determined.

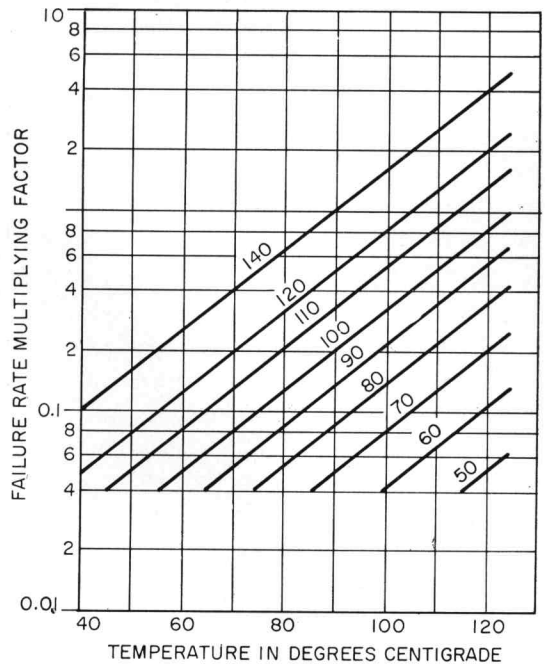
Naturally, the user is not entirely satisfied as he will usually operate the component under more-moderate conditions. Correction factors, which can be applied to the basic rate, thus have to be determined. We touch here on the principle of *accelerated tests* that either allow a smaller number of samples to be handled or the testing time to be shortened. This is of added interest, since the user cannot always wait for one or several years to pass before knowing the quality of a product he must use

now. We will consider this fundamental problem at greater length subsequently.

2.4.1 *Example 2*

It is nevertheless useful to cite an example from the very-important Parts Specification Management for Reliability Report [3]. It concerns paper or paper-polyester tubular capacitors for operation from 200 to 600 direct volts. Four classes of reliability are set: *M*, *N*, *O*, and *P*, corresponding, respectively, to basic failure rates of 3, 1, 0.1, and 0.01 percent per 1000 hours, defined with a confidence level of 90 percent for a temperature of 125 degrees centigrade and with nominal voltage applied. The manufacturer is supposed to classify his products in this manner after due verification. Figure 1 gives the correction factor to apply to the basic rate for other temperatures and other voltages. These correction factors are the result of an accumulation of observations carried out under various operating conditions and, in the present case, of very-long experience with this type of capacitor. It can be noted, for example, that if a capacitor of the type in question operates under its nominal voltage at only 90 degrees centigrade, the probable failure rate is then only 1/5 of the basic rate.

Figure 1—Correction factors for basic failure rates of paper or paper-polyester tubular capacitors for operation between 200 and 600 direct volts. The percentage of rated voltage is indicated for the various curves.



The very principle of accelerated tests continues to provoke a great deal of discussion, mainly among users who doubt the validity of the extrapolations involved. Although their contention may be valid, we have, nonetheless, little better to propose. Above all, a view that seems to be even more fundamental applies to this example: the failure rates thus predicted include both sudden failures and wear-out failures. The definition of the latter is to a large extent arbitrary because, for example, the lower limit agreed to for the insulation resistance cannot be the same for all applications. We are, therefore, now going to envisage a clearer distinction between these two types of failures.

3. Separation of Failures Due to Aging

It has (provisionally) been considered immaterial whether 500 components are tested for 10 000 hours or 5000 for 1000 hours; this implicitly assumes that λ is constant for such durations, in other words, that neither aging nor infant mortality should intervene in any appreciable manner.

There are ways of limiting the effects of infant mortality, either by preaging the components or by eliminating defective units during the final adjustment of an electronic equipment. This is not the case for wear-out failures, which occur late, so that the above hypothesis is only approximately suitable for short- and medium-term applications.

On the other hand, [1] states, "in the majority of cases, the components of an electronic computer operate in the wear-out phase." It

is thus even more strongly the same for large equipments, present or future, and in particular for the telephone exchanges of tomorrow. The reliable operation of these equipments is no longer determined by the failure rate λ established for the period when such rates are relatively constant, but by higher rates that incorporate wear-out failures.

Another aspect of the question has to be noted here. The failures that occur during the period when λ is constant have a random character. They are generally *total* (short circuit, open circuit, breakdown, et cetera) in the sense that the component is suddenly unable to fulfill *any* function. Such failures are definitely characterized and can be counted without too much ambiguity.

Wear-out, on the contrary, is generally characterized by a progressive alteration in the characteristics of the device (isolation resistance, current gain, et cetera). There is in general no a priori criteria for declaring the component out of use. It depends on the application and the choice of a criterion is essentially arbitrary.

The point of view that these two modes of failure are, in fact, two aspects of a single phenomenon has been expressed. But to avoid deviating from our subject matter through such somewhat theoretical considerations, we will limit ourselves to the phenomenological aspects, that is, there are two main types of failures, one sudden and the other gradual. The user needs information about both.

It is evidently not a question of furnishing such predictions for all the characteristics of various components. On the basis of experience and of common sense, we will limit ourselves to those that are fundamentally representative of the component, such as the capacitance of a capacitor or, particularly sensitive to aging, such as the current gain and I_{CBO} of a transistor.

Moreover, since wear-out phenomena do not intervene until after a time lapse, ideally as long as possible, it is generally necessary to resort to *accelerated* aging procedures to draw conclusions within a reasonable time. A few examples will demonstrate that the prediction of aging phenomena does not present, at least in theory, the same difficulties as the determination of sudden failure rates.

4. Definition of a Failure

4.1 Empirical Methods

A first group of methods of defining a failure follows in Section 4.1.1 and is based on the assumption that, despite what has been said, wear-out failures can be defined without being excessively arbitrary. In a second group (Section 4.1.2), it is assumed that such a decision is left to the user. But the decision is not rigidly governed by the nature of the component. In the following example, which concerns relays, notably, the criterion is based on the observation of the first fault. It might just as well have been left to the user to fix the criterion, on the basis of the evolution of the contact resistance, for instance.

4.1.1 *Cases Where a Failure Criterion May Be Defined*

4.1.1.1 *Example 3*

This example concerns sealed-contact relays. For precision applications, it is convenient and legitimate to consider that a relay has failed as soon as the first fault occurs. In addition, this example happens to be characterized by the discovery of a simple law for the appearance of wear-out failures.

Finally, as a first approximation, it may be accepted that the life of a relay may be expressed as the number of successful operations, rather than in hours. This allows a simple accelerated test consisting of operating the relay at a rapid pace on a given load. Strictly speaking, the operating speed is not totally immaterial for it can modify the heating of the contacts, among other things. But let us disregard this refinement.

Figure 2 represents the results of a life test on 44 sealed-contact *H80* relays of Standard Elektrik Lorenz. The test continued for 200 million operations at a frequency of 35 per second, and the relays switched 100 milliamperes of direct current in a resistive load of 500 ohms.

The diagram represents the percentage of survivals as a function of the number of operations. The ordinate scale is such that a normal, Laplacian, or Gaussian distribution as a function of the abscissa is represented by a straight line. This choice results from the fact that over the wear-out period the distribution of failures as a function of the logarithm of the number of operations is, precisely, Gaussian. The distribution is then qualified as log-normal. The bottom of the curve (dotted) corresponds to a constant λ of 1.4×10^{-9} failure per component operation and a statistical mean life $\bar{t} = 7 \times 10^{8}$ operations (if the process were continued indefinitely and without change).

A slightly more rigorous statistical analysis can be carried out by means of Table 1 and by substituting component operations for component-hours of test. There are 3 failures during the first 50 million operations per relay, that is, $50 \times 44 \times 10^{6} = 2.2 \times 10^{9}$ component-operations roughly. Since $k = 6.7$ for $P = 0.90$, the results are as follows

$$\lambda_{max} \text{ (90 percent confidence level)} = k/nh \approx 3 \times 10^{-9}.$$

This underlines the difference between an observed failure rate and a maximum probable failure rate with a given degree of confidence.

In summary, there is a superposition of two mechanisms: up to 50-million operations, the failures are random only and appear at the constant rates defined above. Beyond 50-million operations, wear-out failures preponderate, with a life duration of 130-million operations, defined as the moment when *50 percent of the devices are out of service.** The standard deviation is equal to the interval that separates the points for 84 percent and for 50 percent, which, taking into account the logarithmic scale in abscissa, are −46 and +86 percent. Naturally, if the law of failures could not be represented by a normal distribution, it would nonetheless be desirable to furnish the value of the standard deviation which, however, would have to be calculated instead of being so simply determined graphically.

The test can be accelerated by other than frequency of operation, such as by the influence of the switched current on life. This has been applied to the case of the *G29* type of sealed relays of the Western Electric Company [4]. We will not deal with this example, which permits derivation of the law for the average life as a function of the switched current.

Three important conclusions can be drawn from example 3.

(A) An ideal reliability test should furnish

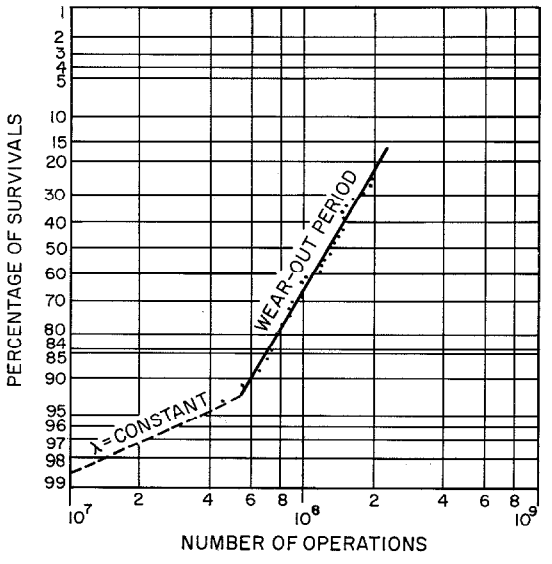


Figure 2—Life test on sealed-contact relays.

* τ is not identical to $\bar{t}$ as defined above except in particular cases, notably the present one of a normal distribution; but not in general.

information about both random failures and wear-out failures.

(B) Without entering into sophisticated analysis, it can be admitted on the basis of the good alignment of the values observed that probable life predictions have a high degree of confidence, even for as modest a sample as 44 devices.

(C) It can also be said that preliminary conclusions can be formulated long before the end of the test, for example, after 80 or 90 million operations in this case. This permits the test to be made either more economical or of shorter duration.

4.1.1.2 *Example 4*

In the preceding example, knowledge of the failure-distribution law permitted the use of an appropriate graph without any further consideration and with all the related advantages, in particular representation by a straight line. If a simple classical law cannot be uncovered a priori, the Weibull distribution [5], which is extremely general in application, can be used. It is expressed by a survival law in the form

$$R = \frac{N}{N_0} = \exp\left[- \frac{(x - \gamma)^\beta}{\alpha} \right]. \quad (8)$$

(Defined for $x \geq \gamma$.)

x is the parameter of evolution, such as time, number of cycles, et cetera, and α, β, and γ are constants.

Stange [6] and Kao [7], among others, have shown the extensive potentialities of such a function.

For $\beta = 1$ and $\gamma = 0$, it is identified with an exponential distribution (see (2A)). For $\beta = 3.25$, it represents a normal distribution,

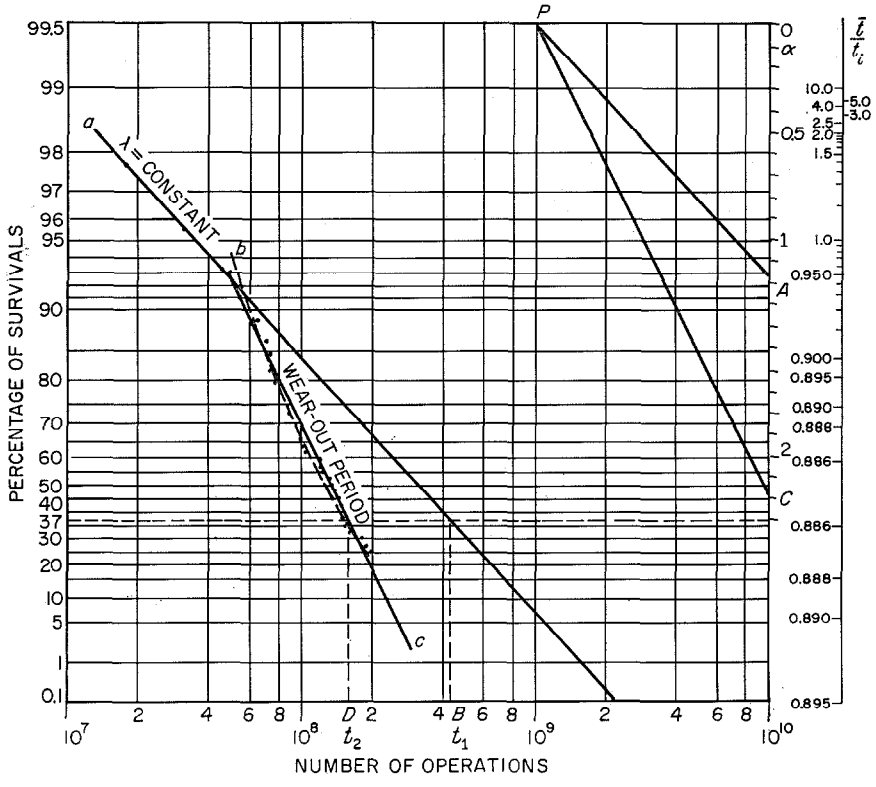


Figure 3—Weibull-Stange graph of life test in sealed-contact relays.

with an acceptable numerical approximation. Another advantage results from the relation

$$\log \ln (1/R) = -\log \alpha + \beta \log (x - \gamma). \quad (9)$$

It can be seen that on a graph where the ordinate scale is graded in (log log) and the abscissae in (log), a straight line with a slope proportional to β results. Making observations of R over a long period, we can thus expect to find two approximately straight lines intersecting at a certain angle, one corresponding to the quasi-exponential period (random failures) with $\beta \approx 1$, the other to the aging period with $\beta > 1$.

The use of Weibull's functions and graphs in itself implies precautions and refinements that are too lengthy to outline in this article. We will simply give an example, in the form of Figure 3, which is given the Stange presentation. To clearly delineate the differences with example 3, the same case of sealed-contact relays has been chosen.

It can be observed that the first 3 failures correspond to points situated exactly on a straight line, the slope of which corresponds to $\beta = 1.18$ (point A); that is to say, the hypothesis of an exponential law ($\beta = 1$) made in example 3 appears justified. A graphical treatment, on which we will not dwell otherwise, gives from points A and B the following statistical mean life duration

$$\bar{t} = 4.2 \times 10^8 \times 0.945 \approx 4 \times 10^8 \text{ operations.}$$

Analysis of the second part of curve bc is a little more delicate. $\beta = 2.2$ (point C) is found, corresponding to the fact that effectively a log-normal distribution is more nearly approximated than a normal distribution. The points C and D give an average life duration of

$$\bar{t}_2 = 1.6 \times 10^8 \times 0.885 \approx 1.4 \times 10^8 \text{ operations.}$$

In both cases, we find about the same orders of magnitude as in example 3. Strictly speaking, the points of the segment bc are better represented by the dotted curve than by the straight line bc. But this can lead to considerations that are unnecessary here. It may be said briefly that the Weibull-Stange representation appears to be, in the present case, better than the preceding one for segment ab and not as good for segment bc. The fact that it directly provides the statistical mean life is not a major advantage since the knowledge corresponding to x percent survivals is generally more useful, and can be seen directly on the diagram.

In sum, the Weibull-Stange representation is always useful as a last resort. A perfect representation is not at all basic theoretically, but only important in practice. In fact, the experimental points are generally scattered around the average curve. An improper choice of graph form thus can result in appreciable differences in the statistical parameters: average value (compare examples 3 and 4), standard deviation, et cetera. The resulting errors in the extrapolations that follow can become rather considerable.

4.1.2 *Cases Where No Failure Criterion Is Fixed*

Whereas in the preceding two examples, the failure (end-of-life) criterion, although arbitrary, was nonetheless very precise, it is now assumed that it is preferable for the user to determine that criterion according to the application envisaged. One then is limited to furnishing statistical data on the evolution of characteristics with time.

4.1.2.1 *Example 5*

The method will be applied below to dry tantalum capacitors. Because of the complexity of the problem, we will limit ourselves here to the characteristic that evolves most rapidly, leakage current. The capacitors were manufactured by Standard Elektrik Lorenz and were tested at the Laboratoire Central de Télécommunications by what one might call an accelerated test, although in this particular

example the operating conditions were not more severe than the limits authorized by the manufacturer. It would probably have been advantageous if they had been.

Capacitors rated at 4 microfarads and 35 volts were subdivided into 6 groups of 50 capacitors each and subjected to three temperatures (55, 70, and 85 degrees centigrade) and two voltages (17.5 and 35 volts, with a protective resistance of 100 ohms) for a maximum time of 8000 hours.

Figure 4 represents the evolution of the leakage current I_l for the group that was subjected to nominal voltage at an ambient temperature of 85 degrees centigrade. Curve A shows the leakage-current distribution at the start. The distribution turned out not to be strictly log-normal. By distorting the ordinate scale in an appropriate and empirical fashion, a nearly rectilinear average curve was obtained. This artifice is perhaps surprising, but it is nonetheless legitimate. The exact mathematical law representing the distribution is not of great importance, providing a single straight line is obtained; it is only a question of convenience and graphical precision. The thing to be carefully noted is that the scale selected is satisfying only between 2 and 98 percent. Consequently, the diagram has been truncated accordingly.

The straight lines B, C, and D show the situation at 1000, 4000, and 8000 hours, respectively. To avoid complicating the graph, the corresponding experimental points

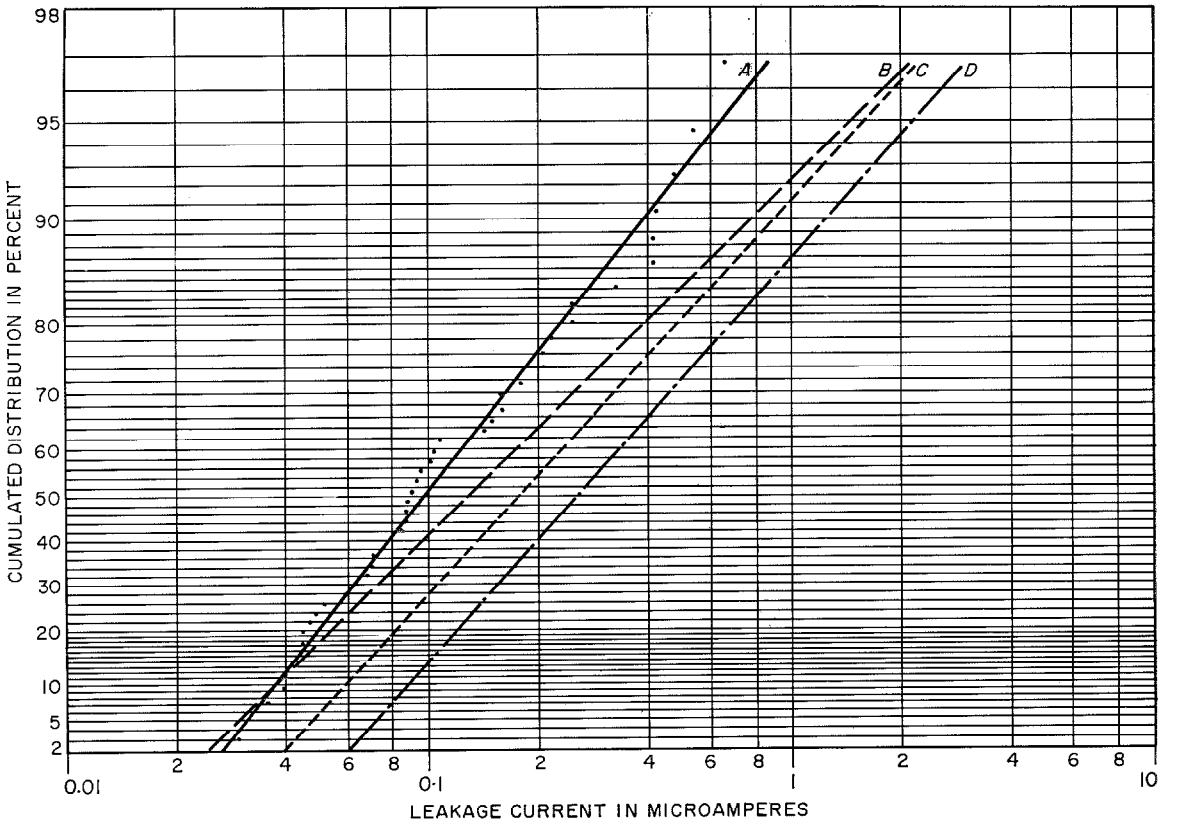


Figure 4—Reliability test on tantalum capacitors. Curve A is at the start of the test and curves B, C, and D, are after 1000, 4000, and 8000 hours, respectively.

were not plotted but their alignment was satisfactory.

Taking this diagram as a start, the evolution with time of any reference point whatever, for example the median value of 50 percent, can be studied. This is done on Figure 5, curve A. It is clear that an extrapolation up to 100 000 hours, although rather uncertain, can be essayed. The shaded area is purely figurative, but it can be considered unlikely that the median value of I_l will exceed 10 microamperes before 100 000 hours. Certainly an experimental point at 15 000 hours, even 20 000 hours, will greatly diminish the uncertainty when it is obtained.

Curve B shows the evolution of the 95th percentile. The uncertainty is a little bit greater because of the rather peculiar behavior of the capacitors up to the 1000th hour (curve B of Figure 4). It can be estimated, however, that 5 percent of the capacitors will have a leakage current exceeding 10 microamperes at around 30 000 hours.

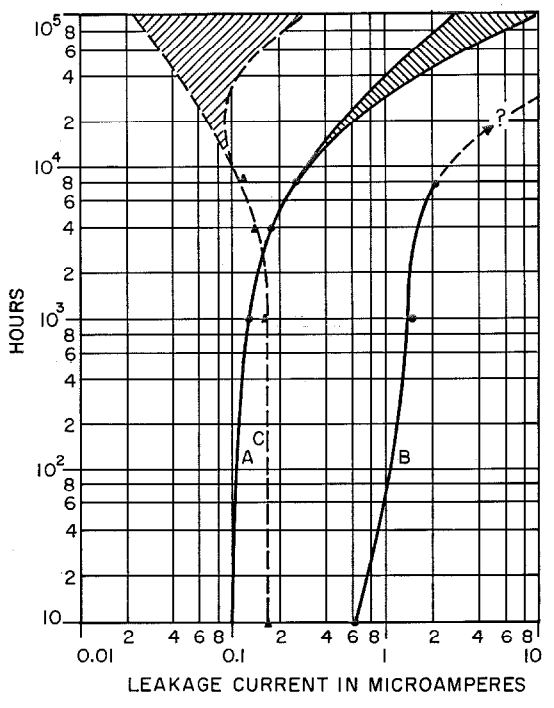


Figure 5—Extrapolations of reliability data on tantalum capacitors.

The other groups in the test can be similarly treated. Thus, curve C of Figure 5 represents the evolution of the I_l median value for the group submitted to nominal voltage at 55 degrees centigrade. Here, there is even an "improvement" up to 8000 hours. The extrapolation becomes very uncertain, and the shaded area is again purely figurative. But in all probability, there will not be any pronounced deterioration before 100 000 hours at 55 degrees centigrade.

4.2 Utilization of a Working Hypothesis or Model

In the preceding examples, the predictions were the result of the direct analysis of experimental observations which could either be represented by a classical mathematical law (Gaussian, Weibull, et cetera) or essentially exploited by graph. It goes without saying that if the mathematical *form* of the law of failures or wear-out is known, it is easier to devise the tests with a view to determining this law *quantitatively* (verifying it at the same time). When aging processes are basically physico-chemical in nature, it can be admitted that their speed with time v as a function of the absolute temperature T obeys Arrhenius' law.

$$v = A \exp [- (B/T)] \qquad (10)$$

where A and B are constants.

Passing to logarithms

$$\ln v = \ln A - (B/T). \qquad (11)$$

Thus, $\log v$ is a linear function of the inverse of the absolute temperature, which makes extrapolation easy. It can be admitted that the acceleration of a phenomenon is equivalent to changing the time scale, and therefore that the acceleration affects equally the random failures and the wear-out failures. However, this is not self-evident.

More-elaborate models have been proposed, for instance Eyring's. But the Arrhenius model is usually sufficient. It has been applied with success in numerous cases, notably that of semiconductors. As in Section 4.1, we should have made a distinction between cases where a failure criterion is fixed and those where this is not done. To limit the length of this article, we will consider only the case where a certain criterion is arbitrarily fixed.

The work of D. S. Peck of the Bell Telephone Laboratories is authoritative in this domain. Many other examples are described in the technical literature on the subject [8]. The example below is the result of work carried out at the Transistor Division of Standard Telephones and Cables.

4.2.1 *Example 6*

Germanium alloyed transistors were divided into groups, each subjected to a different temperature, that is 135, 115, 100, and 75 degrees centigrade. It had been previously recognized that the parameter most sensitive to degradation was the I_{CBO}, and the value of 10 microamperes was fixed as the failure criterion. The failures versus time were then plotted on a graph analogous to that used for the relays in Example 3. The alignment of points indicated a good fit with a log-normal distribution. It does not seem necessary to reproduce this diagram. After the examples already described, one can imagine that it is easy to find the moment when, for indicated temperatures, the total percentage of failures reaches whatever value one may fix: 10, 50, or 90 percent, for instance. We will limit ourselves to recording these values in Figure 6, where the abscissa scale is graded in $1/T$ and the ordinates are logarithmic.

Taking the inevitable graphical errors into account, the agreement between the theory and the actual experiment can be considered very satisfactory.

For the 50th percentile, the straight-line extrapolation is based on the points for 135 and 115 degrees centigrade. The point at 100 degrees is a short distance from this straight line, while the point at 75 degrees is too low. (As indicated, the point at 100 degrees was not observed but was calculated from other observations.) The situation is similar for the 10th percentile.

The discrepancy affecting the points at 75 degrees has since been explained. The corresponding transistors were from an earlier production lot and were not quite so good as the others. The test at 75 degrees was rerun with transistors from the same period as the

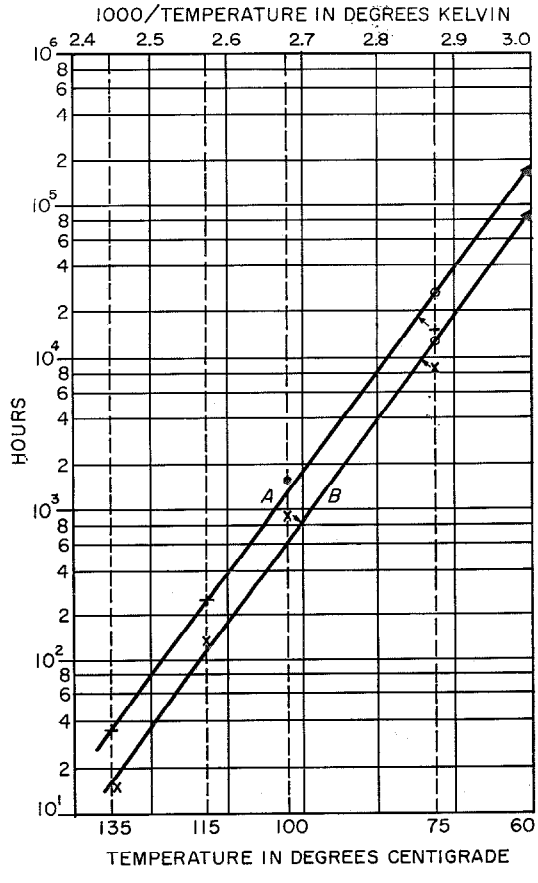


Figure 6—Reliability test on transistors based on the Arrhenius model. Curve A is for the 50th percentile and B is for the 10th percentile. The crosses are observed points, the dots are computed, and the circles are extrapolated.

others and the points now tend to be correctly aligned.

A relatively moderate extrapolation reveals that at 60 degrees centigrade (without dissipation of power in the transistors), 10 percent of the transistors will fail after 90 000 hours (10 years) and 50 percent after 180 000 hours (20 years). Whatever uncertainties may affect this result, it is extremely valuable, and better methods to establish such predictions are hardly conceivable.

Arrhenius' model can be applied with equal success to film resistors for variations in resistance as a function of surface temperature but not for failures as above.

A more-general character can be obtained by utilizing an effective temperature given by

$$T_{\text{effective}} = T_{\text{ambient}} + aP_d \qquad (12)$$

a being a constant and P_d the power dissipated in each device. There are therefore considerable potentialities in the Arrhenius model.

5. Synthesis and Conclusions

5.1 Recapitulation of Difficulties

The principal limiting factors involved in directly determining *very-small* failure rates can be summarized as follows.

(**A**) The hypothesis λ = a constant, which can be applied to only a limited period in the life of a component.

(**B**) The defining of both random failures and wear-out failures by a single parameter.

(**C**) The high number of component-hours of testing.

(**D**) The representativeness of a sample.

(**E**) The validity of extrapolations in accelerated tests.

(**F**) The often-disregarded multiplicity of factors, such as; humidity, shock, vibration, radiation, and human errors.

5.2 Proposed Solutions

We have considered at length the importance of separating random failures and wear-out failures. Whereas accelerated tests on moderate quantities of samples give sufficiently precise results insofar as the law of aging is concerned, difficulties are encountered in their use to determine very-small rates of *random* failures.

5.2.1 *Utilization of Sequential Tests*

We quote a passage from a very-significant editorial [9].

"Testing components to see how many in a group will fail within a certain time or testing components to destruction and determining why they failed is no longer satisfactory in many cases. The number of failures that will occur in a group of high-reliability components within a practical test period may be so insignificant that no conclusive evidence is found."

Reliability tests being long and *destructive*, it is not conceivable to test 100 percent of production; only sample quantities may be tested. This raises the question of the representativeness of any sample.

Thus, in view of the orders of magnitude we are now approaching, anxiety exists as to whether a sampling really represents the average quality of the production considered and also whether test results have not been distorted by a banal accident, clumsiness, or error, even in the absence of any suspicion of misrepresentation.

Difficulties (**C**) and (**D**) above, would seem to be in the process of solution along the general lines indicated in the reliability report already cited [3]. The cost and length of reliability tests will limit their application to average levels, which eventually might be the subject of guarantees on the part of the manufacturer. Beyond, reliability will no longer be determined by qualification or acceptance tests. A new spirit, based on

cooperation and reciprocal confidence between manufacturers and users, has to be developed. The manufacturers should set up a highly developed organization for quality assurance and control, not only to improve quality, but particularly to achieve uniformity of quality.

Then the determination of failure rates will assume an entirely different aspect. The manufacturer will be obliged to maintain accurate records of failures observed in all circumstances: manufacturing controls, tests for customers, et cetera. In accordance with the idea expressed above, only random and total failures should be taken into consideration. There are, then, several ways of statistically cumulating the corresponding information by putting the tests end to end, so to speak. This permits the accumulation of a large number of component-hours of tests, covering an extended period of the particular production.

5.2.1.1 *Example 7*

A first method is based on (6) and values from Table 1 for a confidence level of 90 percent, transferred to Figure 7. To demonstrate the principle, a series of tests carried out at Laboratoire Central de Télécommunications on tantalum capacitors from Standard Elektrik Lorenz has been analyzed. The temperature was not the same in all cases, but it is estimated that 80 degrees centigrade

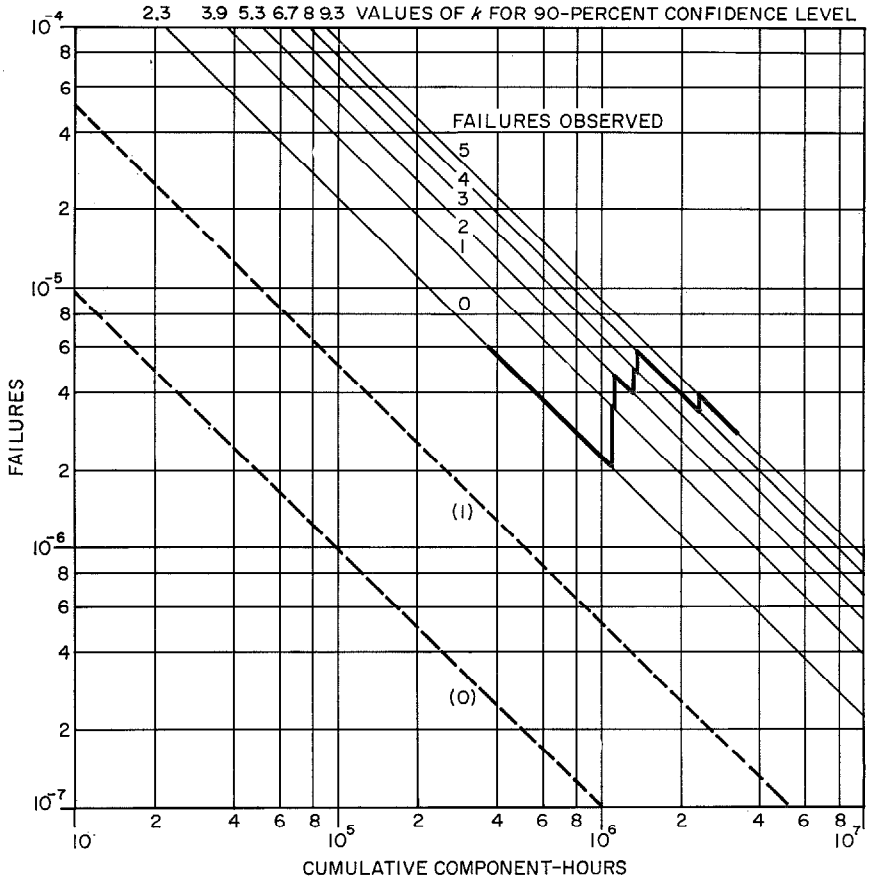


Figure 7—Cumulative reliability tests on tantalum capacitors based on (6) and Table 1.

TABLE 2
CUMULATIVE DATA FROM TESTS ON TANTALUM CAPACITORS

Test	Quantity	Failures	Accumulated Component-hours	
			Per Test	Total
1	10	0	220 000	220 000
2	30	0	720 000	940 000
3	25	1st at 7000 hours 2nd at 8000 hours end of test	175 000 +24 000 +92 000	1 115 000 1 139 000 1 231 000
4	50	1st at 2500 hours 2nd at 3000 hours end of test	125 000 +24 500 +336 000	1 356 000 1 380 500 1 716 500
5	50	0	500 000	2 216 500
6	50	1st at 3000 hours end of test	150 000 +343 000	2 366 500 2 709 500
7	50	0	500 000	3 209 500

represents a fair average. The results that were calculated are given in Table 2 and are recorded in Figure 7. Now, after a cumulation of 3 200 000 component-hours of test, it appears, with a 90-percent confidence level, that λ is less than 3 *total* failures per million component-hours at 80 degrees centigrade, which is very satisfactory. It is useful to compare this with the gross rate observed, 1.6 failures per million component-hours. The advantage of this presentation is that it furnishes an upper limit for λ, *with no preconceived idea* as to the value expected (similarly, the dotted straight lines give the lower limit). A disadvantage is, perhaps, compression of the divisions through the use of logarithmic scales.

In general, cumulative tests of this nature, called sequential, are handled in a different and much-more-elaborate fashion from the statistical point of view. The principles have been particularly developed by B. Epstein and M. Sobel [10], and A. Wald [11]. In these methods, a compromise is sought between the risk for the user of overestimation of life and the risk for the manufacturer of underestimation of life.

Complete analysis of these methods is known to statisticians but involves more-elaborate probability concepts, such as the chi-square distribution, and is decidedly outside the framework of this article.

5.2.1.2 *Example 8*

Poisson's law is a somewhat-different very-simple presentation that also derived from (4), (6), and Table 1. Figure 8 is the corresponding diagram. It is rendered more useful by grading the abscissae with $1/\lambda_1$ as unity, λ_1 being the target rate *to be verified*, which, being fixed a priori, is contrary to the preceding method. The significance of the curves is indicated in Figure 8.

Let us take example 7 for comparison and determine the situation for a failure rate of $\lambda_1 = 2 \times 10^{-6}$. As abscissa unit we then have $1/\lambda_1 = 0.5 \times 10^6$ component-hours. Let us record in ordinates the failures listed in Table 2 at the moments they occurred. This results in a stepped curve.

Let us arbitrarily assume that the production will be considered qualified for $\lambda = 2 \times 10^{-6}$

if this value is assured with a 95-percent confidence level (90 percent would be just as admissible). This would be the case if the stepped curve did not exceed curve A, a possible result which however is far from obtained in this case. At the same time, and just as arbitrarily, it can be considered that the production is disqualified for $\lambda = 10^{-6}$ if this rate is reached with only a 5-percent probability; thus corresponds to curve E and thus the production is effectively disqualified for $\lambda = 10^{-6}$. If $\lambda = 3 \times 10^{-6}$ is considered, curve F is the result, with a probability of 95 percent. It can be seen that the production has just been qualified after 1.1×10^6 component-hours. It is prudent to await confirmation by a new intersection, which seems likely to occur, without this being absolutely certain. A 90-percent probability that $\lambda = 3 \times 10^6$ is shown by curve G and is in agreement with Figure 7. Summing up, having agreed on acceptance conditions for a rate λ_1 and rejection conditions for a rate λ_2 (below λ_1, in general), intersection with the corresponding curve is awaited before a decision can be taken. This presentation, as can be seen, is somewhat different from the usual presentation for sequential tests, but it is simple in principle and in application. The reference curves can be traced on transparent sheets, which then can be placed over the diagram containing the stepped curve.

The report [3] already cited recommends that by means of sequential tests all manufacturers can achieve failure-rate qualifications that normally should improve with time. It is possible, however, to go backwards, through

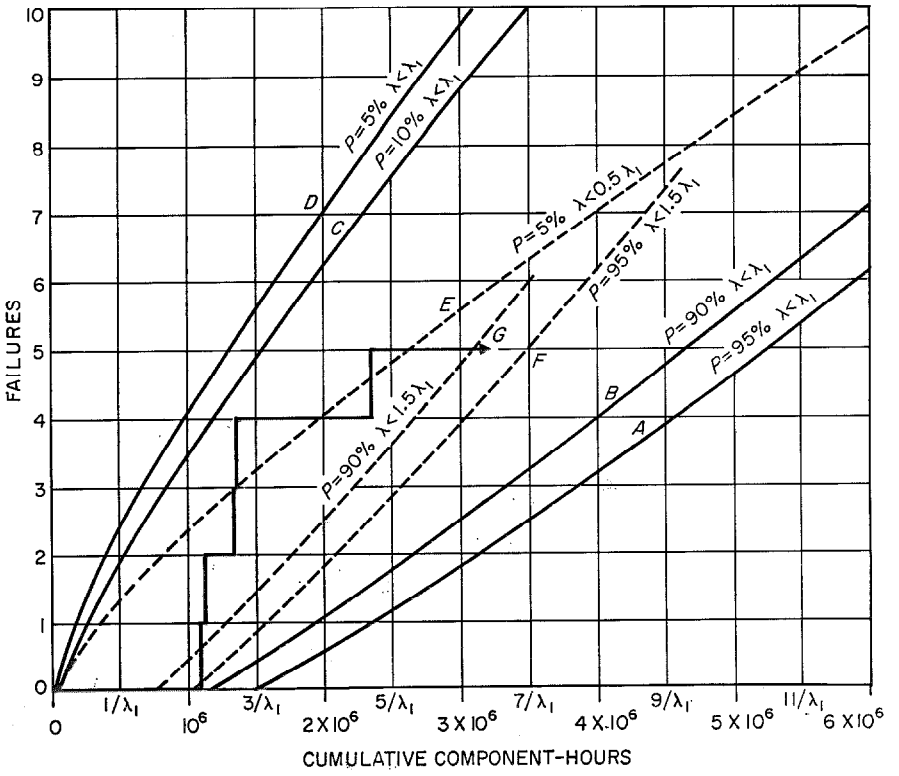


Figure 8—Scquential tests of tantalum capacitors. λ_1 is the target failure rate. In the first example, $1/\lambda_1$ equals 5×10^5 component-hours. The lettered curves are for various failure rates λ for the confidence levels P indicated.

the temporary loss of a qualification. A considerable discussion would be required on this subject.

An extension of this policy would lead to the pooling of confirmed information originating with one or several manufacturers and one or several users. There is a marked tendency in this direction in the form of setting up veritable testing "pools." The Electronic Component Reliability Center is an example of such. This group is sponsored by 13 important firms and it redistributes analyzed and synthetized information. But the publication of this information is private and it cannot be cited here.

Thus the tendency is toward a national or international professional reliability organization, which does not eliminate the need for independent testing by each company.

5.2.2 *Other Problems to Be Solved*

There are two further difficulties that we have not yet considered. One is the validity of the extrapolations connected with accelerated tests. The conditions for acceleration may introduce a profound alteration in aging mechanisms instead of a real acceleration. This is in part plausible and it is a physicist rather than a statistician who must penetrate the depths of knowledge of aging phenomena. It is obviously senseless to carry out accelerated tests at the melting point of solder.

On the other hand, an intransigent attitude does not appear justifiable. It is clear that aging is the result of a number of superimposed processes, which are different functions of exterior conditions, notably temperature. Thus a particular mechanism A can be preponderant at low temperature and another B preponderant at high temperature. It would seem that the essential thing is for the transition to be *continuous* and, if possible, *monotonous* (in the mathematical sense of the word). Then, it is always possible to discover a theoretical or empirical quantitative law for the whole phenomenon that will permit the extrapolation.

The latter considerations relate to the more-imponderable parameters such as: humidity, vibrations, human errors, et cetera. Unfortunately, no correlation has yet been established with a sufficient precision between the results of climatic tests and the behavior of a given component under normal conditions. This represents a tremendous gap, and attempts are currently being made to fill it. The same is true for shock, vibration, et cetera. For radiation, deterioration can in principle be accelerated by exposure to a very-dense radiation flux.

There remains as the second difficulty the question of human errors, which are very difficult to evaluate. We can only state that if a single failure will appreciably affect the results of a test, a test fault involving such a failure is no longer tolerable. The tests and notably the measurements and recording of parameters should be entirely automatic, as should the numerical calculations that follow.

6. General Conclusion

Today, the reliability of electronic components has reached such high levels that its determination is becoming more and more difficult. The obstacles to be met transcend the possibilities of even large companies and it is only by a pooling of efforts that a solution may be found. If this summary of the situation seems pessimistic, it is for the user's sake. The latter should constantly keep in mind the numerous uncertainties that influence measurements in this domain and consequently should provide for safety margins in accordance with the importance of the envisaged application. In particular, he will have to use with circumspection the failure-rate values published without sufficient information on the conditions under which they have been established.

7. Acknowledgments

The author wishes to thank the various testing laboratories of the ITT Europe companies that furnished an abundance of examples of methods and experimental results. His gratitude is addressed to the following persons in particular: Messrs. P. Morain and J. Simonet (relay testing at Le Matériel Téléphonique); M. B. Derjavitch (capacitor testing at Laboratoire Central de Télécommunications); Dr. J. M. Groocock (Transistor Division of Standard Telephones and Cables, London); and Dr. W. Ackmann (of Standard Elektrik Lorenz, concerning the Weibull-Stange methods).

8. Symbols

A, B = coefficients of Arrhenius' law
a = numerical coefficient
h = number of hours of test
$k = nh \times \lambda_1$
m = number of failures observed
N = number of survivals at time t
N_0 = number of components in the tested lot
n = number of samples
P = confidence level
P_d = dissipated power
R = proportion of survivals at time t
S = proportion of failures at time t
T = Absolute temperature
t = time
$\bar{t}(\bar{t}_1, \bar{t}_2 \cdots)$ = statistical mean life
v = speed of reaction
x = unknown or unspecified variable
α, β, γ = coefficients of Weibull's function
λ = failure rate (in general)
λ_0 = observed failure rate
$\lambda_1(\lambda_2 \cdots)$ = specified (expected) failure rate
τ = life at 50-percent survivals.

9. References

1. S. R. Calabro, "Reliability Principles and Practices," McGraw-Hill Book Company, New York; 1962.

2. J. H. Scrivner and J. R. Willey, "Proving Long Term Reliability," *Electronic Industries*, volume 21, pages 102–106; May 1962.

3. "Parts Specification Management for Reliability," Superintendent of Documents, United States Government Printing Office, Washington 25, District of Columbia.

4. O. M. Hovgaard and W. J. Fontana, "A Miniature Versatile Switching Capsule," Proceedings of the Electronic Components Conference in Philadelphia, 1959; pages 116–121.

5. W. Weibull, "Statistical Distribution Function of Wide Applicability," *Journal of Applied Mechanics*, volume 18, pages 293–297; September 1951.

6. K. Stange, *Milteilungsblatt für Mathematische Statistik*, Jahrgang 7, pages 113–151; 1955.

7. John H. K. Kao, "A Summary of Some New Techniques on Failure Analysis," 6th National Symposium on Reliability and Quality Control in Electronics, Washington, 1960.

8. J. E. Shwop and H. J. Sullivan, "Semiconductor Reliability," Engineering Publishers, Elizabeth, New Jersey; 1960.

9. Editorial, "Failure and Reliability," *Electronics*, volume 35, page 3; 12 October 1962.

10. B. Epstein and M. Sobel, "Sequential Life Tests in the Experimental Case," *Annals of Mathematical Statistics*, pages 82–93; March 1955.

11. A. Wald, "Sequential Analysis," J. Wiley and Sons, New York; 1947.

Charles A. Meuleau was born in Paris, France, in 1908. He graduated from the Ecole Supérieure de Physique et Chimie and also received the License ès Sciences degree from the University of Paris in 1928.

From 1931 to 1938, he was in the chemical industry. Previously, from 1929 to 1931, and after 1938, he served in the International System at the Paris laboratories, now known as Laboratoire Central de Télécommunications. From 1950 to 1958, he was head of the semiconductor products department. He is now with the Technical Office of ITT Europe headquarters, where he performs a coordinating function related to components. He has worked on plastics, varnishes, natural and artificial piezoelectric crystals, thermistors, semiconductors, and electronic measurements of physical properties.

Mr. Meuleau collaborated with Professor Goudet in preparing the book "Les Semiconducteurs—Diodes, Transistors et autres Applications."

Microelectronic Devices

C. P. SANDBANK
Standard Telephones and Cables Limited; Footscray, Sidcup, Kent, England

1. Introduction

The large electronic systems of to-day use circuits in which a particular assembly of components may appear hundreds of times. Designers are therefore thinking in terms of integrated circuit units, which may consist simply of two or three components in one package or which could be more-complex units equivalent, for example, to one or more stages of a shift register.

Although the impetus for the development of microelectronic components came mainly from the need for smallness, this is often not the main advantage of such devices. In this paper various microelectronic devices are described, and some of the reasons for using particular integrated circuits in place of the separate components are discussed.

2. Types of Microelectronic Devices

The term "microelectronic device" is used here when referring to a miniature component in general, and "integrated circuit" is used to refer to the particular case of microelectronic device having a number of interconnected components intended to perform a specific circuit function. Integrated circuits would normally have only the input, output, and supply terminations externally accessible.

The terms "passive components" (resistors, capacitors, et cetera) and "active components" (transistors, diodes, et cetera) are used to denote the circuit function and not the construction. Thus, a reverse-biased diode intended to act as a capacitor in a semiconductor circuit would be referred to as a passive component.

The various microelectronic devices can be divided into the groups shown in Table 1.

The first group, Integrated-Circuit Modules, is subdivided into standard components assembled in sealed units and into special components, such as multiple devices containing a number of active devices in only one package. Components in a special mounting intended for automatic assembly and interconnection would fall into the latter subdivision.

The second group, Thin-Film Circuits, consists mainly of two types, one having evaporated or sputtered resistors and capacitors with active elements that are subsequently mounted on the circuit separately. The other type, having evaporated active components, is still in a relatively early stage of development.

The third group, Semiconductor Circuits, has all the active and passive components fabricated on the semiconductor by processes similar to those used to manufacture transistors. The significance of the common-substrate connection and isolated-substrate connection will be discussed later.

The groups listed in Table 1 are intended as a general guide only and do not give a rigid

TABLE 1
MICROELECTRONIC DEVICES

Integrated-Circuit Modules		Thin-Film Circuits	Semiconductor Circuits
Modules with Standard Components	Modules with Specially Made Components	Evaporated Passive Components with Transistors and Diodes in Miniature Encapsulations Evaporated Active and Passive Components	Common-Substrate-Connection Circuits Isolated-Substrate Circuits
Transistor-Resistor Tunnel-Diode-Resistor Circuits Complex Logic and Linear Circuits	Multiple Devices Complex Micromodules Tunnel-Diode-Resistor Modules		

classification. There is considerable overlapping. For example, some modules could equally be classed as a multiple device or as a semiconductor circuit.

Figure 1—Matched pair of germanium tunnel diodes coded type *JK60A*.

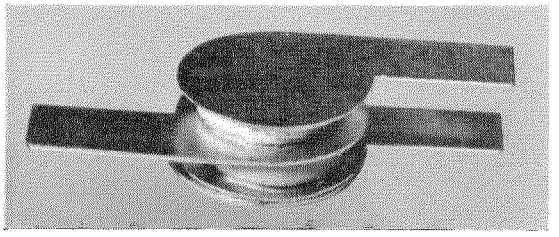


Figure 2—Dimensions of *JK60A*.

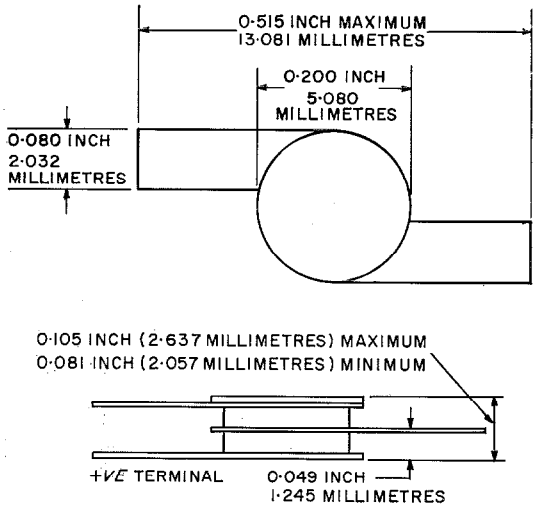
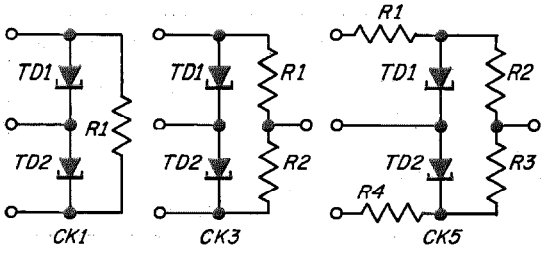


Figure 3—Circuit configurations for typical logic elements based on a Goto pair.



3. Integrated Tunnel-Diode Circuits

The tunnel diode is perhaps the most immediately rewarding device to use in integrated circuits. With relatively simple modules it is possible to achieve quite spectacular improvements over the performance of circuits using lumped components. To obtain the best performance with tunnel-diode circuits it is necessary to maintain close matching of certain device characteristics and to keep the circuit series-inductance values as low as possible. These conditions can best be met with integrated modules.

3.1 Tunnel Logic Elements

The Goto pair, two tunnel diodes connected in series, forms the basis of many high-speed tunnel-diode switching circuits. In this circuit the speed depends on close matching of the peak currents of the two diodes and also matching of the junction capacitances. By mounting two diodes in a microminiature dual package of the type shown in Figure 1, it is possible to obtain an over-all series inductance of less than 2 nanohenries. The dimensions of the package are shown in Figure 2. The method of manufacture enables the peak currents to be matched to within 2·5 per cent within the range of 5 milliamperes ± 10 per cent and the junction capacitances matched to less than 10 picofarads about a nominal 25 picofarads, on the *JK60A* device illustrated.

For maximum speed in a bi-stable circuit it is necessary to integrate some of the other components with the matched pair to keep the over-all circuit inductance low. Circuits of the type shown in Figure 3 will operate with diodes having the characteristics given above, at clock frequencies of several hundred megacycles per second, provided the inductance of each resistor and its leads is less than 1 nanohenry. An example of one of these circuits is shown at the extreme right of Figure 4.

To obtain unidirectionality of the logic flow it is usual to use such devices in synchronous

systems. With 3-phase systems, decision rates of over 1 million per second are feasible.

The size of the circuit modules is approximately 3/8 by 5/16 by 3/16 inch (9·5 by 8 by 5 millimetres) plus tags. This is small enough to allow high packing density thus reducing propagation delay in the assembled system. Since tunnel diodes are low-level devices, there should be no difficulty with excessive heat dissipation due to high packing density.

The use of the technique of making fast integrated circuits for strip-line systems is of course not limited to tunnel-diode devices. Any system using ultra-high-speed majority- or minority-carrier devices must have very-low-inductance circuits and low-delay connections of the type obtained in the tunnel logic elements.

3.2 Tunnel Memory Elements

The fastest known storage systems can be made with tunnel-diode memory-circuit elements of the type shown in Figure 5. The speed of such stores, which have access times of around 100 nanoseconds, is not, in this instance, limited by the switching speed of the tunnel elements. It is therefore not so necessary to use low-inductance components. The speed is largely limited by the rewrite cycle, which in turn depends on the propagation delay along the read direction. Another very-important factor in the operation of tunnel-diode stores is the sum of the tolerances of the individual components that go to make up each bit.

The main reasons for integrating are therefore:—

(A) To provide suitably shaped small elements that can be stacked close together to reduce the propagation delay in the read direction.

(B) To allow the most-economical combination of component tolerances to be chosen. This can best be achieved during manufacture.

(C) To reduce the over-all size, weight, and complexity of the store. A memory element

Figure 4—Tunnel-diode integrated circuits.

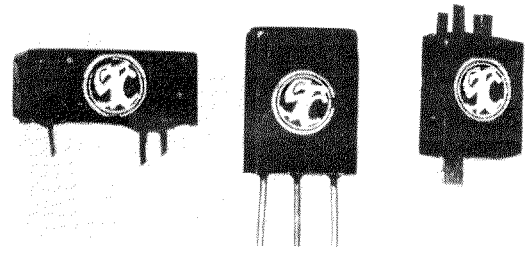


Figure 5—Tunnel-diode memory-circuit configurations.

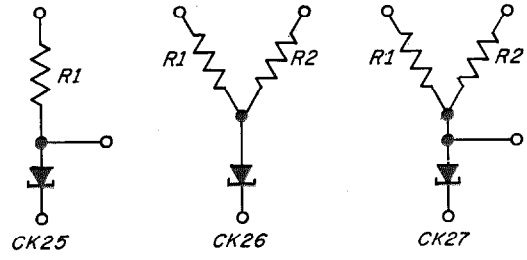
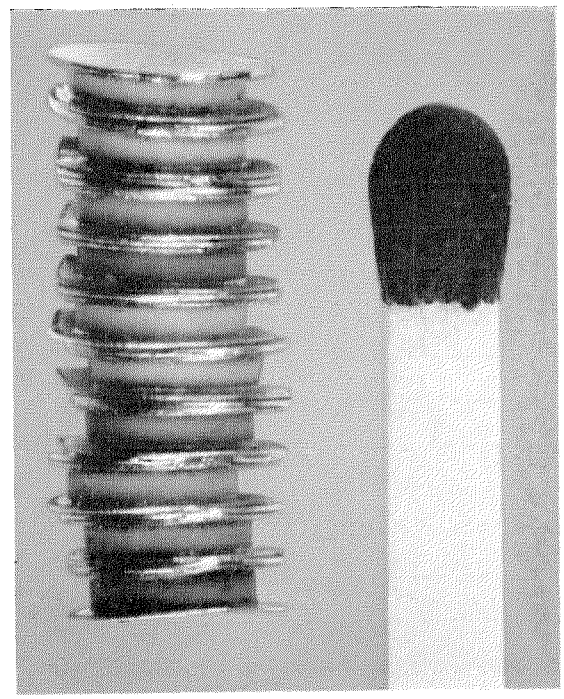


Figure 6—*CK50* before encapsulation.



type *CK26* is illustrated at the centre of Figure 4, having the circuit configuration shown in Figure 5. The modules have been made flat, so that they can be stacked closely together. A stack of 32 words in a 1000-bit store would be approximately 5 inches (127 millimetres) deep by 1 foot (0·305 metre) long.

3.3 Tunnel-Diode Counting Elements

By using a chain of tunnel diodes connected in series, it is possible to perform specialized counting operations. For example, *CK50* has 10 tunnel diodes with peak currents graded so that they switch sequentially in response to input pulses. This unit can be used as the basis of a simple decade scaler circuit. Units of this type may be readily integrated. Figure 6 shows the assembly of 10 tunnel diodes, before encapsulation. The finished device with terminals brought out from the ends of the chain and also from the penultimate diode, for counting and resetting, can be seen at the left of Figure 4.

Figure 7 is a reproduction of the voltage-current characteristic of the device. The peak currents of the diodes range from 7 to 13 milliamperes in equal steps. The maximum valley current is 2 milliamperes.

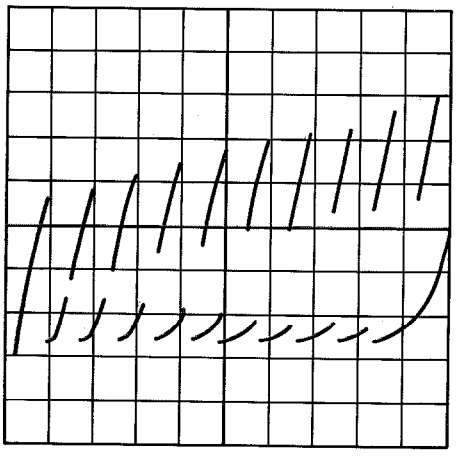


Figure 7—Typical current-voltage characteristic of *CK50*.

4. Multiple Devices

Multiple devices are packages containing more than one active device. Since the mounting and sealing processes for most low-level transistor and diode elements are common, it is advantageous to encapsulate widely used configurations of active devices together in one package. This represents a compromise between using separate components and complete semiconductor circuits. Although multiple devices do not have the advantage of completely integrated circuits, they are more flexible in application, and fit in readily with existing systems using established circuits. Multiple devices can be made either by mounting and connecting the individual elements on suitably isolated areas of the encapsulation or, alternatively, by using semiconductor circuit techniques so that more than one element may be produced in the semiconductor crystal mounted on the header. Multiple devices are frequently mounted in standard transistor outline packages having a suitably increased number of leads, depending on the complexity of the device. A commonly used header has the *TO5* outline with 8 leads, but even smaller headers with the *TO18* outline have been made with up to 10 leads.

The three configurations shown in Figure 8 have been chosen to illustrate some of the reasons for using multiple devices. The technique can, of course, be extended to more-complex configurations. For example, packages containing all the active components for diode-transistor logic gates have been made.

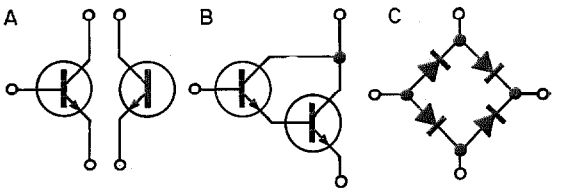


Figure 8—Circuit configurations for multiple devices. *A* is a double transistor, *B* is a Darlington amplifier, and *C* is a diode bridge.

4.1 Double Transistors

Transistors are frequently used in pairs. Where matching is required this can best be carried out during manufacture, and ensured by mounting the devices in a dual package. In direct-current amplifiers using balanced pairs it is particularly important to keep the operating temperatures of the two transistor elements as close as possible. This can best be achieved if the devices are in close thermal contact.

Figure 9 is a photograph of the double transistor type *BFY20*. This is in a *TO5* outline header. The two transistor elements are alloyed to the two blocks set in the glass base. This ensures that the elements are in good thermal contact whilst remaining electrically isolated.

4.2 Darlington Amplifier

The configuration of the Darlington amplifier is such that the terminations are the same as for a transistor. A standard transistor header can therefore be used, as shown in Figure 10. With devices of this type current gains of several thousand can be obtained. The advantages of using a multiple-device Darlington amplifier in place of two separate transistors are:—

(**A**) Increased reliability due to the over-all reduction in the number of connections between dissimilar alloys.

(**B**) Reduction in size.

(**C**) Saving in the cost of the additional transistor case and encapsulation process.

Functionally the Darlington amplifier is a simple semiconductor circuit, because the two transistors are terminated individually as are the transistors or diodes shown at *A* and *C* in Figure 8. The method described here is the multiple-substrate technique of making semiconductor circuits. Another method of making this amplifier is described in Section 6.1.

4.3 Diode Networks

Two or more diodes are used together in many logic and linear circuits. Among the examples are sets of diodes with common *p* or *n* regions. These can be mounted together on one chip, giving increased reliability and

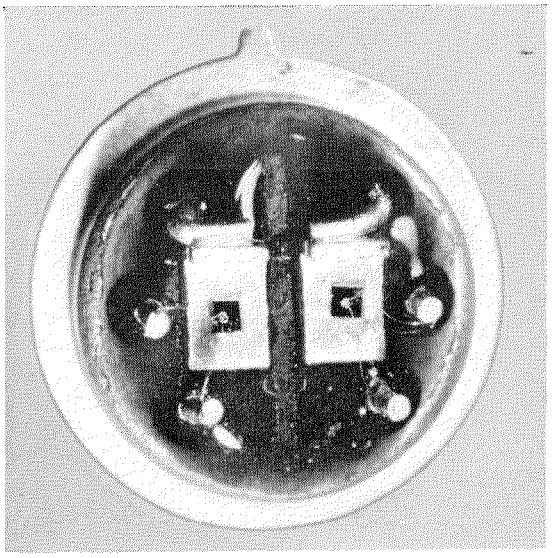


Figure 9—Double transistors in a single mounting type *BFY20*.

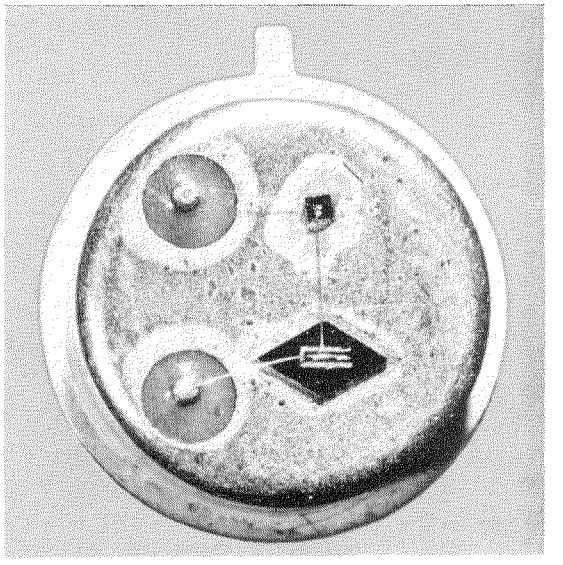


Figure 10—Darlington amplifier.

uniformity of characteristics. Diode bridge and ring networks can be pre-connected in small packages ready for direct use in the circuit boards.

5. Thin-Film Circuits

The two most-widely used techniques for fabricating the passive components on glass or ceramic substrate are:—

(A) Sputtered tantalum for resistors with an oxide of tantalum as dielectric for the capacitors.

(B) Evaporated nichrome for resistors with an oxide of silicon as dielectric for the capacitors.

Both methods enable components to be fabricated with tolerances better than 5 per cent. The nichrome resistors have a lower temperature coefficient than the tantalum resistors. The capacitors using silicon oxide have better power factors than those with tantalum oxide but they occupy a greater area of the substrate. Thin-film circuits and semiconductor circuits are regarded by many users as directly competitive. Since both techniques are in relatively early stages of commercial use it is too soon to predict their relative application or importance when they are in widespread use. However, certain fundamental technical differences suggest that where large electronic systems are considered, these two types of microelectronic devices are in fact not competitive, but compatible. This will be discussed later, when the characteristics of semiconductor circuits have been described.

5.1 Active Devices for Thin-Film Circuits

Experiments are in progress in many research laboratories on the development of evaporated active devices. Because of the difficulty of evaporating suitably doped single-crystal semiconductors, it has not yet been possible to make satisfactory transistors or diodes in this manner. The most promising experiments have been concerned with the manufacture of evaporated field-effect devices and tunnel triodes based on dielectrics such as cadmium-sulphide, which are readily evaporated with aligned crystallites. However, most practical thin-film circuits use separate transistors and diodes wired into the circuit.

One of the difficulties of producing active devices for thin-film circuits has been to provide a small package meeting the encapsulation requirements for the reliable performance of germanium alloy or silicon mesa transistors.

The advent of the planar transistor has created new possibilities in extending the scope and increasing the packing density of thin-film circuits.

The inherent surface passivation of the planar transistor suggests that it may be possible to use the devices mounted directly in thin-film circuits. Thin-film circuits are normally given some protection from the atmosphere in service, and the passivation of the transistor could protect the device until the circuit is finally sealed. An example of a thin-film circuit with unencapsulated transistors soldered directly to the contact areas is shown in Figure 11. However, this involves the handling, storing, and testing of minute devices having leads sometimes less than 0·0005 inch (0·0127 millimetre) in diameter. Where this is im-

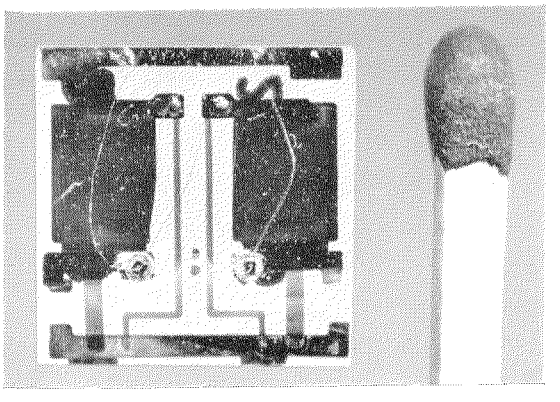


Figure 11—Thin-film circuit with directly wired unencapsulated planar transistors.

practicable, it is possible to take advantage of the fact that the planar device does not need to be encapsulated in a hermetically sealed envelope, but can be sealed in a solid package that can be made much smaller. The other important advantage of planar transistors when used with thin-film circuits is that they combine the properties of high speed with high gain at low currents. This enables efficient low-level circuits to be designed so that a high packing density can be used without the risk of overheating the equipment.

5.2 Solid Encapsulated Transistors

Figure 12 is a drawing of the "Miniflake" transistor. This device consists of a flat plate with part of the area occupied by the encapsulated device, the remainder being metallized areas to which the emitter, collector, and base contacts can be made. Such devices overcome the difficulties of handling and mounting the unencapsulated planar transistors; yet the over-all height of Miniflake transistors is less than 0·025 inch (0·635 millimetre).

A double-ended-contact pattern gives increased freedom in circuit layout; it enables the device (which has an insulating lower surface) to be used for crossovers if the crossing lead can be tied to one of the three transistor electrodes, and allows the device to be securely anchored to the circuit by its leads.

Figure 13 shows two Miniflake transistors mounted in a thin-film multivibrator circuit. Note that the transistors lie over the area occupied by the passive components, thus giving the maximum component packing density. Where the contact pattern on the thin-film circuit has been specially laid out to connect with the transistor, the device may be inverted and sweated directly to the circuit without intermediate leads.

6. Semiconductor Circuits

Semiconductor circuits have been made in germanium and also in silicon using mesa or planar construction. Research is being carried out in using the highly insulating properties

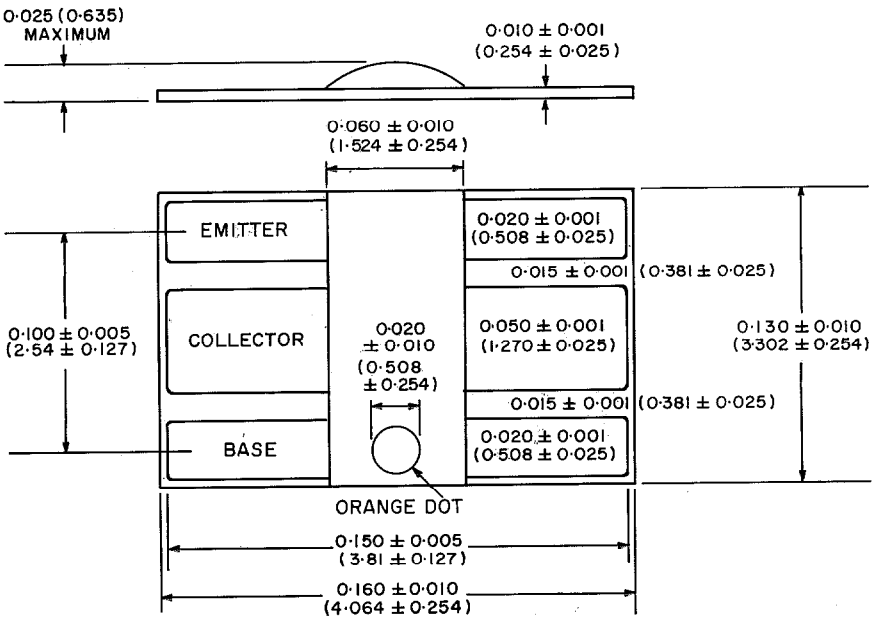
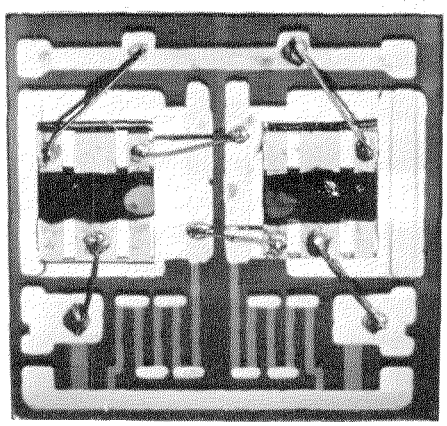


Figure 12—Dimensions in inches (millimetres) of a Miniflake transistor.

of intrinsic gallium arsenide as a substrate for semiconductor circuits. In this paper only the silicon planar technique will be described. By using essentially the same processes as for making transistors, it is possible at the same time to fabricate also diodes, resistors, and capacitors, but not inductors. It is possible to simulate resonant circuits with resistors and capacitors by using series-parallel connection, or by introducing piezoelectric elements into the circuit.

Figure 13—Thin-film circuit with Miniflake transistors.

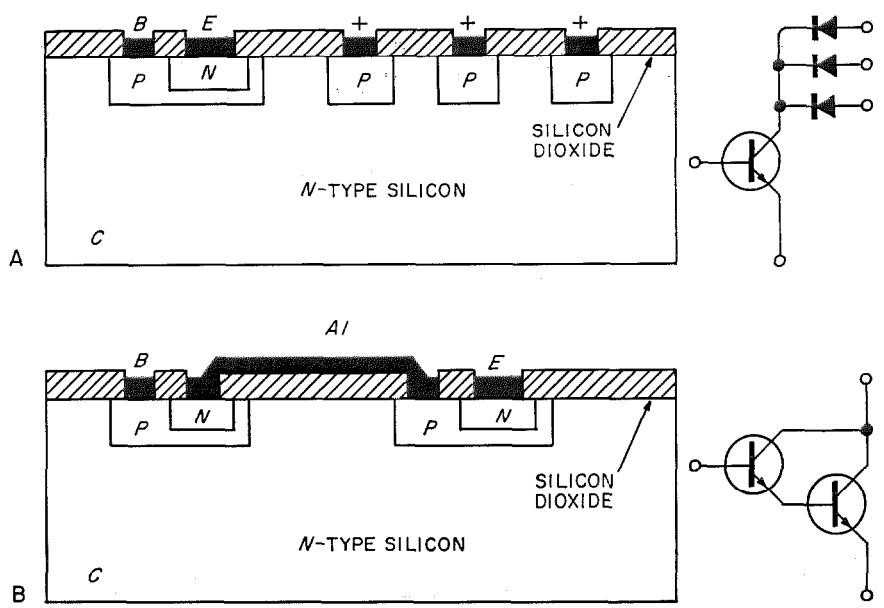


6.1 Semiconductor Circuits with Common-Substrate Connection

Many useful circuit elements can be made with configurations of devices having all n-type or p-type electrodes common. It is thus possible to diffuse the devices in adjacent positions on the silicon, using the standard transistor-fabricating techniques. The semiconducting substrate provides the common electrode. This type of construction is the simplest form of semiconductor circuit.

Two circuits using common substrate connection are shown in Figure 14. A transistor gate with three output diodes is shown at A. Starting with an n-type substrate, the p-type base region of the transistor and the three diodes are diffused at the same time, using the silicon-diode masking techniques of the planar-transistor process. The position and geometry of the devices are controlled by photolithographic techniques to within 0·0001 inch (0·0025 millimetre). Whilst the diodes are protected by a new layer of oxide, the

Figure 14—Semiconductor circuits with common substrate connections. A is a transistor diode gate and B is a Darlington amplifier.



emitter of the transistor is diffused through a window photo-etched in the oxide.

A semiconductor version of the Darlington amplifier can be made using the configuration shown in Figure 14B. This consists essentially of a low-level transistor adjacent to a high-level transistor. The collectors of the two transistors are joined by the common *n*-type substrate. The emitter of the low-level transistor is joined to the base of the high-level transistor by an evaporated aluminium film over the insulating silicon-dioxide layer.

It is possible to introduce passive components into semiconductor circuits with common substrate by using the bulk resistivity and junction capacitance.

Common-substrate-connection semiconductor circuits impose severe limitations on the type of circuit configuration that can be used, but their electrical performance resembles more closely the behaviour of the equivalent conventional circuit than that of the isolated-substrate semiconductor circuit.

6.2 Isolated-Substrate Semiconductor Circuits

6.2.1 *Technology*

To obtain the greatest freedom in circuit design it is necessary to have all the components electrically isolated. This is achieved by introducing an extra *p-n* junction into the semiconductor circuit as shown in Figure 15. Consider first the transistor at Figure 15A. Starting with a *p*-type substrate, for example, the silicon is oxidized as in the first stage of making a planar transistor. A hole is now etched corresponding to the total area to be occupied on the silicon by the transistor. A deep *n*-type diffusion is now carried out through the oxide mask resulting in a *p-n* junction to substrate. If this junction is reverse-biased, then the *n*-type region remains isolated from the substrate. The base region and emitter region are diffused using the standard planar transistor technology.

Contacts are alloyed to the three regions, thus giving access to all the electrodes from the top surface of the slice. A diode is made in a similar way but with only two diffusions. By maintaining the substrate at the most-negative potential of the circuit, all the device-to-substrate junctions will be reverse biased and all the components will remain electrically insulated. The silicon has a protective oxide covering all but the component contact areas so that components can be interconnected by evaporating aluminium leads across the top of the slice as in the case of the common-substrate-connection semiconductor circuit.

Resistors are made by cutting a narrow channel in the oxide layer during the first diffusion and making contact to the ends of the diffused path, as shown in Figure 15B. Capacitors are made by using large-area reverse-biased junctions. The dielectric properties of the silicon dioxide can also be used to form capacitors.

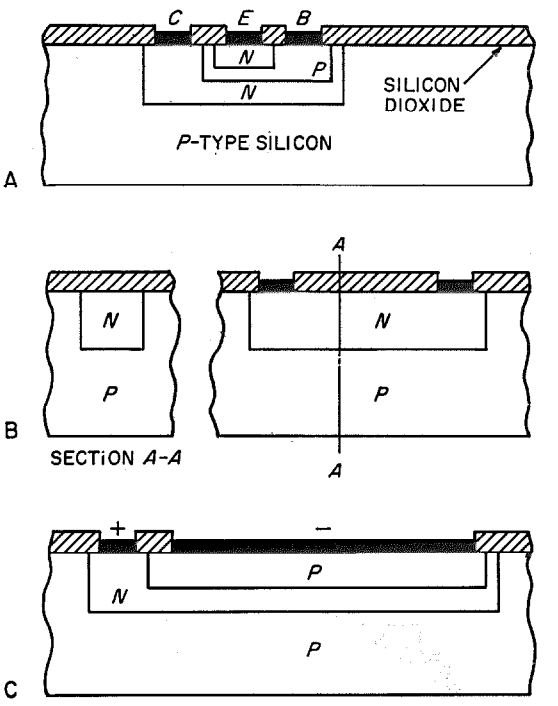


Figure 15—Isolated substrate semiconductor circuit components. *A* is a transistor, *B* is a resistor, and *C* is a capacitor.

Figure 16 shows the oxide patterns at the contact window-etching stage in the manufacture of a device comprising four transistors (centre), two resistors, and two capacitors. The area occupied by these components is approximately 1·5 by 2·5 millimetres.

Figure 17 shows the device with the evaporated-lead pattern mounted on a *TO5* header with thermo-compression bonding of the leads to the external termination of the circuit.

6.2.2 *Reliability*

Semiconductor circuits should provide the most-reliable circuit element because all the components are made at the same time by the minimum number of processes using very-pure materials and highly controlled techniques. By directly connecting the components on the silicon the possibility of failure of interconnecting is also reduced. From results quoted in the United States the reliability of semiconductor circuits equivalent for example to one complete stage of a shift register could be better than 0·01 per cent per 1000 hours. However, with this sort of reliability many devices must be tested and this is of course a much-more-expensive exercise than testing separate components of comparable reliability. Most of the significant results available to date are based on relatively small numbers of early samples made before manufacturers had established the processes. The results quoted for these circuits are one to two orders less reliable than the estimated figure given above. The performance and reliability of semiconductor circuits of the type described in this section will be the subject of a subsequent paper.

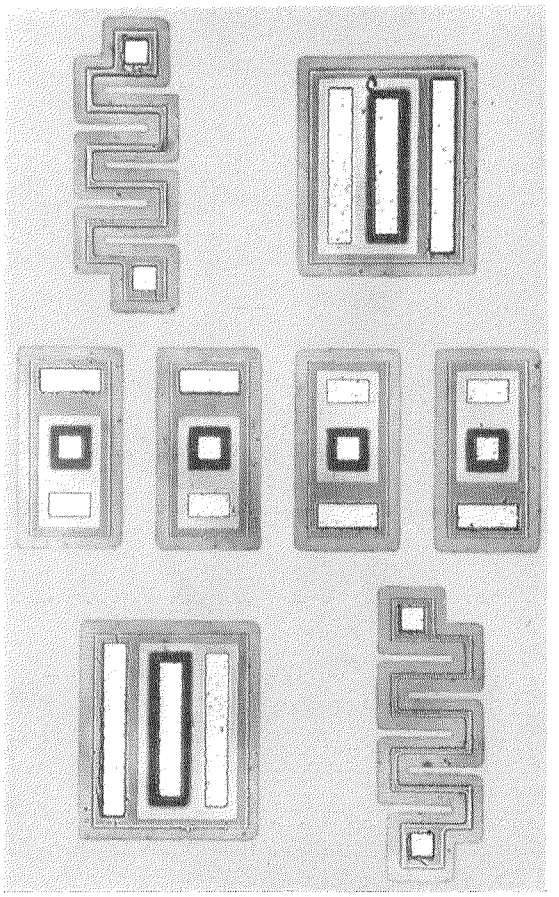


Figure 16—Contact window-etching stage in the manufacture of a semiconductor circuit.

6.2.3 *Packing Density*

The packing density possible in the semiconductor circuit is limited more by heat dissipation considerations than by device size. With low-level logic circuits, very-high packing density is possible. For example, a complete digital integrator using the equivalent of approximately 3000 conventional components per cubic inch has been designed.

6.2.4 *Comparison of Semiconductor Circuits and Thin-Film Circuits*

Returning now to the comparison with thin-film circuits it is clear that on the question of compactness and reliability the semiconductor circuit is to be preferred. The question of cost must be related to the number of devices required. An integrated circuit is obviously more specialized in application than a separate component. The tooling cost of producing a semiconductor circuit (mainly the cost of design and mask manufacture) is very-much higher than the cost of producing the equiva-

lent thin-film circuit. On the other hand, the final cost per circuit should be less for a complete mass-produced semiconductor circuit than for a thin-film circuit plus active devices. Thus on the question of cost it would seem that semiconductor circuits would be preferable where very large numbers are required and thin-film circuits where repetitive circuits are needed in quantities that justify an integrated circuit, but do not justify the cost of a semiconductor circuit. Taking a large digital system for example, leaving out the considerations of reliability and compactness, the shift registers might be made in semiconductor circuits and the read circuits in thin-film circuits.

It is further possible to draw some comparisons in terms of the electrical performance of semiconductor circuits and thin-film circuits. The tolerances on the passive components in semiconductor circuits are coarser than in thin-film circuits. There is also the coupling of components due to the distributed capacitance to substrate in semiconductor circuits whereas the capacitive coupling between components on the thin-film circuits is often no more than on a conventional printed-circuit board. To take these factors into account it may be necessary to carry out considerable re-design before a given circuit function can be performed in the semiconductor circuit, but the corresponding thin-film circuits can be made with essentially the same arrangement of components. Of course both techniques have limitations on the range of component values that can economically be fabricated but higher resistor and capacitor values can be specified in thin-film circuits.

The comparison of electrical performance can be summarized by saying that semiconductor circuits are more suited for applications where the active devices predominate in number and function, whereas thin-film circuits are more suited to applications where passive components predominate. In practice this means that the advantages of the techniques weigh more heavily in favour of digital circuits for semiconductor circuits, and linear circuits for thin-film circuits. This is because in digital circuits the active devices are normally brought from one extreme state to another so that tolerances are of less importance than with linear circuits where equilibrium conditions are maintained by the balance in passive components.

6.2.5 *Combination of Semiconductor Circuits and Thin-Film Circuits*

Naturally such generalizations do not apply in many cases. Calculations on performance or cost can often become meaningless because one essential component in a circuit is not compatible with the technique that is most suitable for all the others. The most-elegant way of solving such problems is to combine the two techniques. This can be done in two ways.

The first method is to use thin-film techniques in conjunction with the semiconductor circuit by using the silicon-dioxide film as substrate for evaporated resistors and as dielectric for capacitors. The circuit shown in Figure 16 has also been made with evaporated resistors in place of the diffused type shown in the

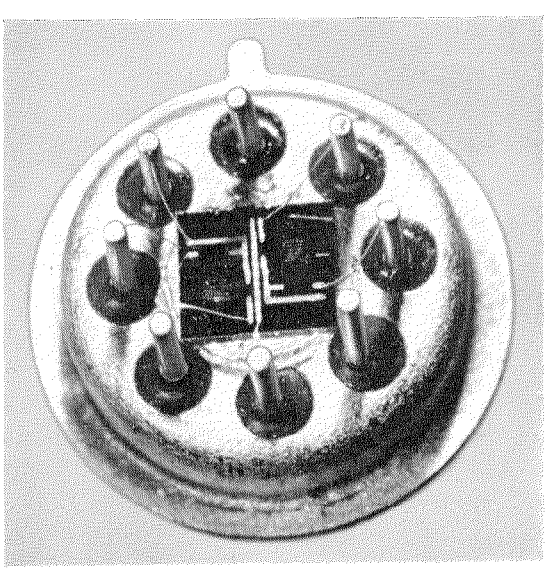


Figure 17—Semiconductor circuit mounted on a *TO5* header.

photograph. These were made with similar sheet resistivity so that in both cases the area occupied was the same. The evaporated resistors are more accurate and have a lower temperature coefficient but the semiconductor circuit made in this way is more expensive because of the extra process involved.

The second method of combining thin-film and semiconductor circuit techniques is to mount the semiconductor circuits on thin-film circuits. Semiconductor circuits having eight active components have been made in a form similar to the Miniflake transistor shown in Figure 12 so that they can be wired directly into thin-film circuits bearing additional associated passive components.

7. Conclusions

Integrated tunnel-diode circuits can be made to perform a variety of logic, memory, and counting functions. In many cases the speeds obtained with integrated tunnel-diode circuits are higher than those possible by any other means. The techniques of making high-speed integrated circuits for strip-line systems will apply to other devices beside tunnel diodes.

Multiple devices can give improved performance particularly in circuits where matched characteristics are important. They can also contribute to the reliability and compactness of circuits.

The use of planar transistors extends the scope of thin-film circuits. The devices can be connected into the circuits in a form that allows high packing density.

Thin-film circuits and semiconductor circuits are not necessarily competitive and will probably both find widespread application within the next few years. Thin-film circuits are more suited for linear and analogue functions whereas semiconductor circuits are at the moment more suited for digital applications. This pattern is likely to remain for some time although there is rapid progress in extending the capabilities in both techniques. In many cases it is advantageous to combine both types of device in a given system or even to combine both techniques in a single device.

Where large quantities are required, semiconductor circuits offer the ultimate advantages in terms of reliability, low cost, and compactness because all the components are made by the same processes as those used for making the transistors.

C. P. Sandbank was born in 1931. He obtained an Honours Degree in physics from London University and the Diploma of the Imperial College in electrical engineering.

In 1953, he joined Standard Telephones and Cables, working as a production engineer, and later as a development engineer. He is now responsible for the development of special semiconductor devices, which include integrated and semiconductor circuits.

Mr. Sandbank is an Associate Member of the Institution of Electrical Engineers.

Silring Mounting of Silicon Power Rectifiers

W. WEISS
Standard Elektrik Lorenz; Stuttgart, Germany

1. Introduction

The high current density in a silicon crystal used as a rectifier generates heat within a small area corresponding to approximately 1 watt per square millimeter of rectifying area. To prevent excessive temperatures that would harm the crystal, this heat is removed by soldering the silicon element to a metal of good electrical and thermal conductivity, usually copper, which is in turn connected thermally to a large-surface cooling unit.

Very-high reverse voltages are also encountered and glass-to-metal and ceramic-to-metal seals have proved to be advantageous for this application.

Permanent protection against deleterious ambient influences, particularly humidity, has required hermetically sealed cases. Manufacturers now use soldering, welding, and cold-welding processes to seal the case.

Basically, the cases of the various commercially available silicon rectifiers are of similar design, the only distinction being the way they are connected to their coolers. Some cases have threaded parts to be screwed to the cooling surfaces; some use flat surfaces for flange-type mounting; and finally pot-shaped cases (automobile rectifiers) with knurled outer walls are pressed into suitably prepared coolers.

Most applications call for combinations of several rectifier cells in bridge circuits, and therefore a space-saving assembly of stacks is very desirable. The Silring mounting provides effectively for compact grouping without sacrificing cooling ability.

2. General Design

Besides having to meet the basic requirements outlined above, it should be easy to assemble rectifier circuits out of individual cells. Unfortunately, this convenience of the well-known assembly principles of the selenium rectifier could not be applied to the conventional case and cooling fin of the silicon rectifier. It was the idea of using the spacing rings between the cooling fins of a stack of selenium rectifiers as cases for the rectifying silicon elements that finally led to the development of Silring.

2.1 Construction

Construction and dimensions of the silicon ring-type rectifier are shown in Figure 1. The case consists of a ceramic ring placed between two nickel-plated copper rings that serve as electric and heat contact areas of the rectifier cell. The silicon crystal is in an eccentric hole in the ceramic ring and soldered to the thicker of the two copper ring plates, connection with the other ring plate being established by a copper wire. The two copper plates are soldered to the partially metallized ceramic ring, thereby hermetically sealing the rectifying element.

The diffused silicon wafer may be soldered in with either polarity. To obtain a better transfer of the heat, however, the crystal element is always soldered to the thicker copper plate which, therefore, should be connected to the cooling plate.

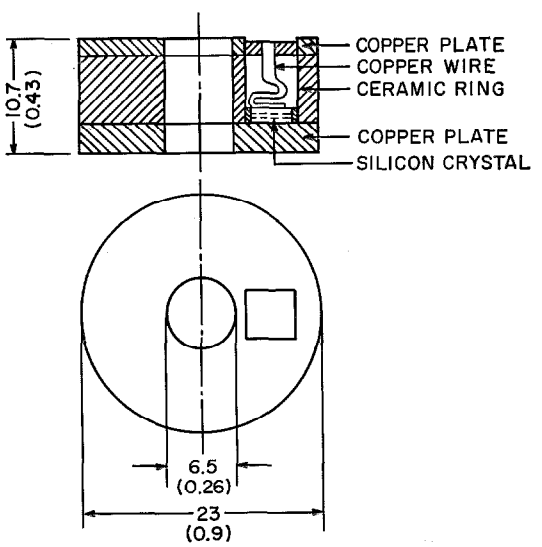
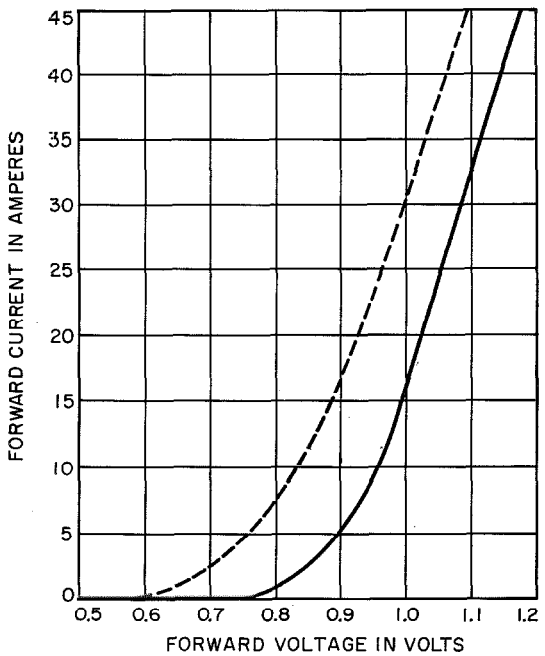


Figure 1—Construction of the Silring silicon rectifier element. Dimensions are given in millimeters (inches).

Figure 2—Mean forward characteristics of the Silring cell at a case temperature of 140 degrees centigrade (broken line) and 25 degrees centigrade (solid line).



2.2 Electrical Characteristics

The physical dimensions of the rectifying element influence the maximum current ratings. Small crystal areas necessitate too many parallel connections for useful ratings while large areas present contact problems and are liable to contain inhomogeneous crystal structures. Large crystal areas also encounter heat transfer difficulties because heat is bound to accumulate at the transition between the case and the cooling plate. This problem can be solved but only through expensive constructional measures.

The maximum current rating of the Silring, based on measurements, is 15 amperes per cell. Figure 2 shows the mean forward characteristics for two different case temperatures. The same ring-type cell may be used for smaller currents. The maximum rated current depends on the dimensions of the aluminum cooling plate. The data are shown in Table 1. The size of the cooling plates ensures that the rectifier case temperature will not exceed the maximum admissible value of 130 degrees centigrade.

The current ratings for an ambient temperature of 35 degrees centigrade, which are normally not given for silicon rectifiers, permit comparison with selenium rectifiers to determine in what particular cases silicon ring-type rectifier stacks are suitable for application under selenium-rectifier operating conditions.

As for conventional silicon rectifier cells, the rated crest working reverse voltage of the Silring depends on the type of crystal used and the manufacturing process employed and may vary between 50 and 600 volts.

3. Silring Stack

3.1 Assembly

The individual ring-type cells, aluminum cooling plates, and connecting tabs are assembled on a steel bolt, which is insulated by a heat-resisting plastic sleeve, and then tightened by a nut. Figure 3 shows a rectifier stack in bridge connection assembled in this manner, with a quarter of the front cooling plate removed to illustrate the arrangement.

Though similar in design and external appearance to the well-known selenium power-rectifier stacks, the ring and the cooling plate

TABLE 1
Ratings of Silring Rectifiers with Convection Cooling

Nominal Current in Amperes		Cooling-Plate Dimensions in Millimeters (Inches)
Ambient Temperatures in Degrees Centigrade		
50	35	
15	17	125 by 125 by 2 (4.9 by 4.9 by 0.08)
10	11.5	75 by 100 by 1 (3 by 3.9 by 0.04)
5	6	50 by 50 by 1 (2 by 2 by 0.04)
2.5	3	Without Cooling Plates

of the silicon ring-type rectifier stacks have distinctly different functions. In selenium rectifier stacks, the plate serves for both rectification and cooling while the ring is used merely as a spacer between the plates. In the silicon ring-type rectifier stack, the plate is used for cooling only and the ring assumes the duty of both rectification and spacing. The aluminum cooling plates are provided with an enamel coating to improve the heat radiation and prevent corrosion.

3.2 Electrical Data

By using suitably dimensioned cooling plates and by connecting the cells in parallel, it is possible to assemble, from only one Silring type, rectifier stacks for rated direct-current outputs of up to 280 amperes with convection cooling and up to 560 amperes with forced air cooling at an air speed of 6 meters (19.7 feet) per second.

Several stacks may be connected in parallel to handle even larger currents, the connecting terminals being designed to permit easy parallel connections.

4. Advantages of Silring

The novel ring-type design of the rectifier case permits an easy assembly of rectifier stacks, with any required internal circuit. The ease of assembly, in turn, permits a high

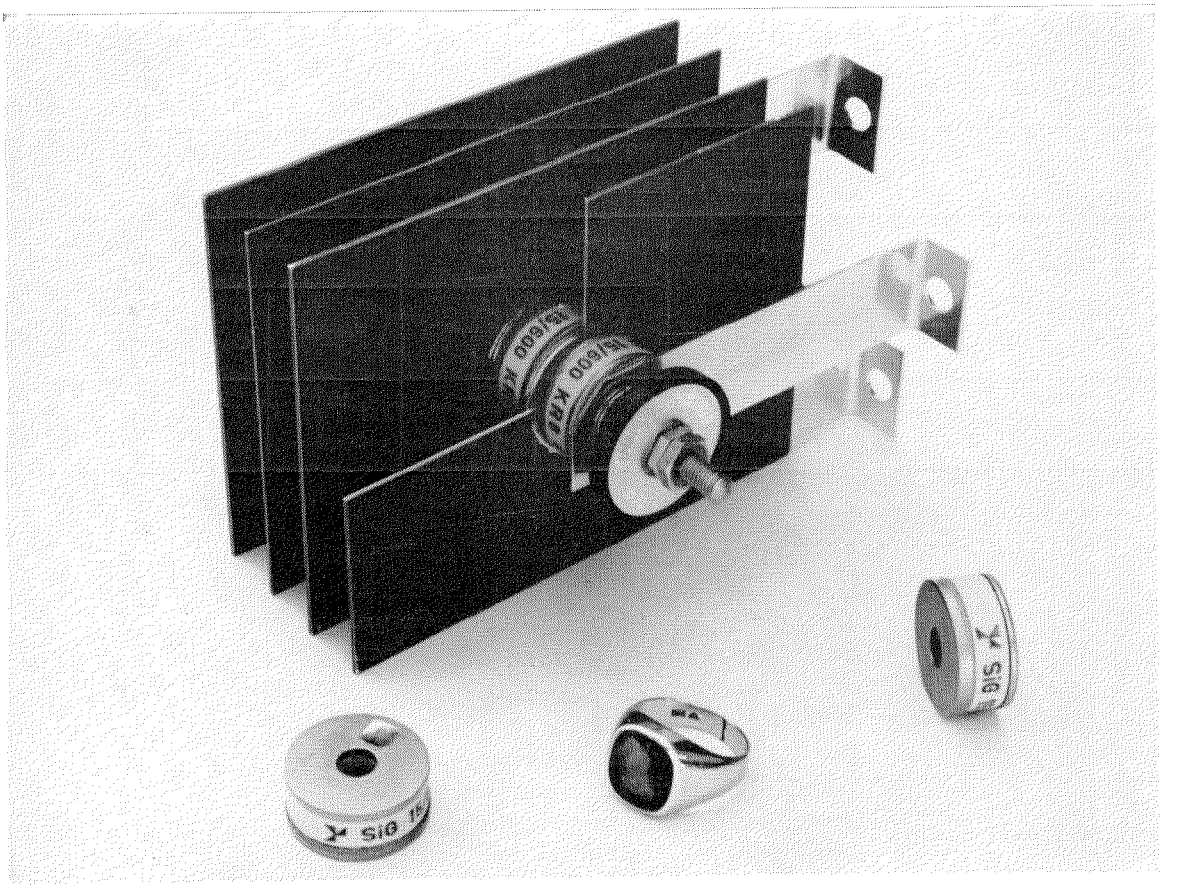


Figure 3—Rectifier stack of four cells in bridge connection with a part of one cooling plate cut away.

degree of standardization for the Silring elements. Even with only one type of Silring cell, rectifier stacks for large currents may be assembled on the building-block principle, necessitating merely different dimensioning of the cooling fins and different internal connections.

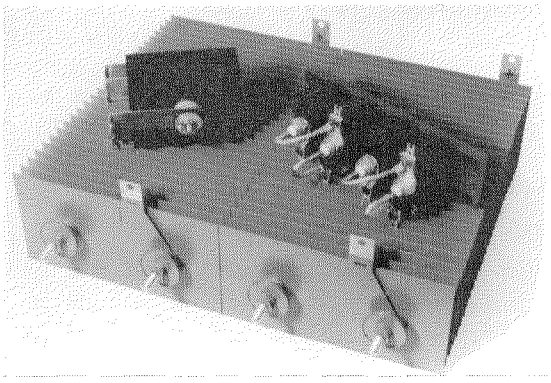


Figure 4—These three rectifiers have approximately the same electrical ratings. Dimensional details are given in Table 2.

The standard elements are dimensioned for reliable, easily controllable, low-cost manufacturing. Spreading of electrical and thermal values is so negligible that no sorting for parallel connection is required.

The flat form of the ring-type case is suitable for spacing of the cooling plates and facilitates mounting and internal circuit connections. Figure 4 shows a Silring, a conventional silicon rectifier, and a selenium unit all having approximately equal electrical ratings. Table 2 gives their dimensions.

During the development work on the silicon ring-type rectifier, close association of the rectifier cell and its cooling system was always envisioned.

TABLE 2
COMPARISON OF SILICON AND GERMANIUM STACKS

Rectifier	Type	Direct-Current Output in Amperes		Maximum Volume		
		Ambient Temperature in Degrees Centigrade		Percent	Cubic Centimeters	Cubic Inches
		35	50			
Selenium	B390/312–24	24	—	100	22 000	1342
Silicon, Conventional	B380/340–20	—	20	8	1 800	110
Silring	B380/340–20	23	20	4	900	55

Werner Weiss was born in Breslau, Germany, on 17 May 1921. He received his doctor's degree in physical chemistry in 1955 from Munich University.

In 1955, he joined the components division of Standard Elektrik Lorenz, where he was engaged in the development of tantalum capacitors, selenium rectifiers, and since 1958 silicon rectifiers. He is now head of the silicon laboratory.

Cause and Prevention of High Reverse Currents in Large-Area High-Voltage Diffused-Silicon Rectifiers

C. F. DRAKE
K. L. ELLINGTON
Standard Telecommunication Laboratories Limited; Harlow, Essex, England

Producers of elcctrical devices are subjected to a continuous pressure, generated by both economic and technological demands, to extend the working range of these components. For silicon devices, there is a three-fold demand; for lower reverse leakage current, for higher working voltages, and always for larger-area devices.

At each stage of technological development it is possible to isolate a particular problem the resolution of which will lead to a further extensive growth. Some years ago, the exact control of the absolute concentration and detailed distribution of significant impurities in silicon devices was the pivotal problem. This control is now a matter of simple routine, and a new restraint on further development has been exposed. The presence of so-called non-significant impurities—any elements other than those of Groups III and V in the periodic table—now determines the quality and yield of the more-sophisticated devices. The problem is not restricted to diffused structures; it arises equally in the production of epitaxially grown silicon devices. The work reported here has been directed towards the identification and classification of deleterious phenomena associated with the contamination of silicon by such non-significant impurities, and towards the development of methods for removing them. The future, however, may well see the effect of the controlled addition of these elements turned to the advantage of the device makers.

1. Reverse Leakage and Breakdown

The reverse current of most silicon diodes is orders of magnitude greater than the theoretical, and the breakdown voltage seldom reaches the design figure. Superimposed on the problem of excess current, is that of the spread of characteristics amongst devices from a production line. The variability is much greater than that attributable to variations in junction depth, surface concentration, or resistivity, and it may become manifest only at elevated temperatures. These effects, of minor importance in components working at less than 200 volts, become increasingly significant as the design voltage is increased, until they dominate the yield and performance of components for use in the 500-to-1000-volt range. Moreover an increase in the junction area often results in a catastrophic decrease in device yield.

1.1 Surface Current

The reverse current may conveniently be considered to consist of two components, one originating at the surface of the junction and the other within the bulk of the material. The surface current may arise from a number of causes. The junction may be partially short-circuited by dirt deposited on the surface, or localized surface breakdown can occur at regions of high electric stress. Both these effects can be readily identified by the sensitivity of the device to further etching and washing, and the effect can be eliminated without difficulty. Current associated with surface states can be distinguished from bulk leakage in two ways. If surface effects predominate, the total reverse current will be sensitive to the surrounding atmosphere and will not decrease proportionally as the area of the diode is reduced. In common with other authors [2–7, 12], we have found that with suitable surface treatment the diode performance is principally determined by bulk effects. The excess reverse current consequent on the bulk properties of the semiconductor will be the subject of the remainder of this discussion.

1.2 Bulk Current

In junctions with narrow space-charge regions operated at low voltages, the voltage-independent diffusion current is important. In high-voltage junctions, thermally excited generation current arising within the space-charge

region, and increasing as a fractional power of the voltage, is the more important at and near room temperature. The dependence of these components of the reverse current on the energy level of the responsible generation centre has been dealt with by Sah, Noyce, and Shockley [8] and by Chevychelov [10]. The effect of specific elements has been considered by many authors [2–5, 7, 9]. At voltages near the design limit, avalanche multiplication produces a rapid increase of current with voltage and the relevant data have been produced by Chynoweth [11].

These sources of reverse current adequately account for the characteristics of ideal or "hard" junctions, but much less attention has hitherto been given to "soft" junctions. The outstanding features of this type of junction are the rapid rise in current with voltage, and the anomalous temperature coefficients. An empirical relationship between current and voltage, $I = KV^n$, where n may vary from 2 to 9, has been generally accepted. Goetzberger and Shockley [12] have suggested that metallic precipitates in the space-charge region are responsible for this behaviour. Sandiford [7] suggested that the carrier emission probability

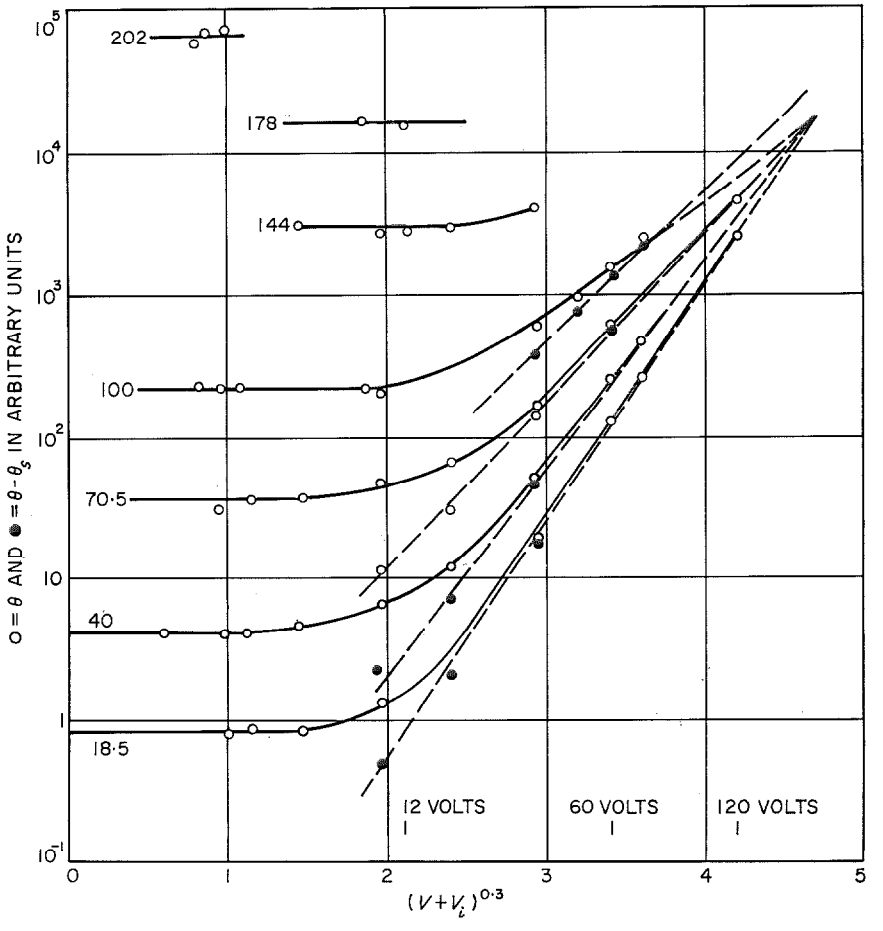


Figure 1—Reverse characteristics of typical soft junction. Gallium was diffused to a depth of 35 microns into 15 ohm-centimetre n-type silicon that had been polished on both faces by a schedule designed to minimize surface damage. The constants for this diode, obtained from this figure and from Figure 2, are given for diode 1 in Table 1. The temperature in degrees centigrade is given for each curve.

is field dependent; whilst Fistul and Orzhevskii [13, 14] have developed the idea of field-dependent emission rates in more detail.

Investigation of the source of the current in soft junctions is clearly a desirable preliminary to removing the cause.

2. Nature of Soft Junctions

Soft junctions were readily obtained by the diffusion of gallium in 10-ohm-centimetre *n* silicon to a depth of 35 microns. The general features of Figures 1 and 2 are typical of such junctions. Soft, epitaxially deposited Zener diodes, despite their lower operating voltage range, exhibit analagous behaviour in the reverse direction at voltages less than breakover.

As the current originates in the space-charge region a normalized current, $\theta = I/(V+V_i)^{1/2\cdot2}$ that is proportional to the carrier-generation rate per unit volume of space-charge region

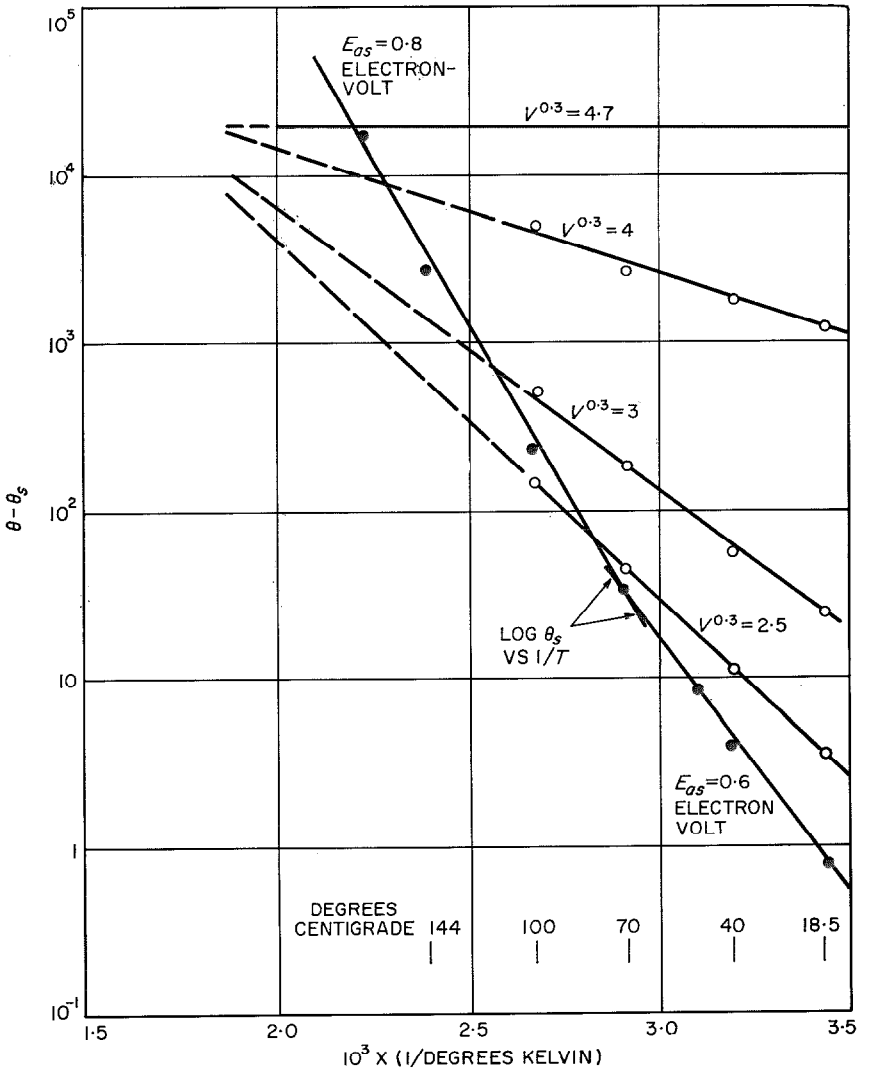


Figure 2—Reverse characteristic of soft junction. The plot of θ_s the saturation current at voltage gives the value of E_{as}.

is used hereafter. The $(V + V_i)^{1/2 \cdot 2}$ dependence of space-charge width on voltage found graphically and by capacitance measurements, agrees well with that found by other workers. The value of V_i, the built-in voltage, appropriate to the temperature, has been used.

2.1 Low-Voltage Region

At very-low reverse voltages (< 2 per cent of V_B) the classical voltage-independent space-charge current θ_s predominates. The temperature dependence is given by $\theta_s = A \exp [- E_{as}/kT]$ where the activation energy, E_{as}, is approximately equal to $E_G - E_{\theta S}$. $E_{\theta S}$ specifies the location of the centre in the band gap and is approximately the depth of the centre from the nearer of the conduction or valence bands. The plot of θ_s against $1/T$ (Figure 2) shows that at temperatures below 70 degrees centigrade the activation energy is 0·6 electron-volt; above this temperature it rises to 0·8 electron-volt. At low voltages, and between 20 and 70 degrees centigrade, the source of carriers is therefore generation centres located at or near the intrinsic level and located within the space-charge regions. It would appear that centres situated about 0·4 electron-volt from a band edge assume the significant role from 80 to 200 degrees centigrade, but the data do not permit an unequivocal distinction between this hypothesis and the alternative possibility that the combined contribution of generation centres within, and a diffusion length beyond, the space-charge region, are responsible for the apparent 0·8 electron-volt activation energy.

2.2 Voltage-Dependent Region

The current in the voltage-dependent region may be approximately represented by $I = KV^n$, but it is difficult to attach any physical significance to this relationship.

Fistul and Orzhevskii [13, 14] have suggested that the effective activation energy of deep generation centres is reduced by the space-charge field to a value $E_{\text{eff}} = E_a - K(E)^{1/2}$ leading to the semi-empirical expression

$$I = A \exp \left[- \frac{(E_a - BV^n)}{kT} \right] \qquad (1)$$

where $n = 0{\cdot}27$ for diffused junctions of this type.

A more-accurate expression can be obtained by replacing I by θ and subtracting the voltage-insensitive space-charge current θ_s when

$$(\theta - \theta_s) = A \exp \left[- \left(\frac{E - BV^{0 \cdot 3}}{kT} \right) \right]. \qquad (2)$$

Preliminary study of the data showed that $n = 0{\cdot}3$ for the diodes investigated. From (2) at constant temperature

$$\frac{\partial \ln (\theta - \theta_s)}{\partial (V^{0 \cdot 3})} = \frac{B}{kT} \qquad (3)$$

and at constant voltage

$$\frac{\partial \ln (\theta - \theta_s)}{\partial \left(\frac{1}{T} \right)} = \frac{- (E_a - BV^{0 \cdot 3})}{k}. \qquad (4)$$

A series of values of $- (E_a - BV^{0 \cdot 3})$ from the straight lines of Figure 2 give self-consistent values of E_a and B. From (3) the slope of the straight lines in Figure 1 should lead to a single value for B. Unfortunately this value of B is dependent on temperature and nowhere agrees with the value obtained from Figure 2.

An additional difficulty is that whilst (4) implies a critical voltage V_c, (such that $BV_c^{0 \cdot 3} = E_a$) at which θ is independent of temperature, the additional experimental observation that the produced straight lines of Figure 2 would converge at a critical temperature T_c at which θ is independent of voltage, is not a consequence of this expression.

The integration of (3) and (4), introducing the experimental observations as integration constants and preserving the observed relationship between V_c, T_c, and $(\theta - \theta_s)_c$, leads to

$$\theta - \theta_s = A \exp \left\{ - \left(\frac{E_a - BV^{0 \cdot 3}}{kT} \right) - CV^{0 \cdot 3} \right\} \qquad (5)$$

or, introducing the experimental constants and re-arranging

$$\theta - \theta_s = \theta_c \exp\left\{-\frac{E_a}{K}\left(1-\frac{V^{0\cdot3}}{V_c^{0\cdot3}}\right)\left(\frac{1}{T}-\frac{1}{T_c}\right)\right\} \quad (6)$$

where the constants

$$V_c^{0\cdot3} = \frac{E_a}{B}$$

$$T_c = \frac{B}{kC}$$

and

$$\theta_c = A \exp\left[-E_a/kT_c\right].$$

The values of E_a and B from Figure 2 are now consistent with the unique value of B from Figure 1. The values of the constants for three diodes, prepared from silicon from different sources and of different degrees of softness, are given in Table 1.

2.3 Discussion

It is clear that a new form of behaviour intervened in practice before the values V_c and T_c were reached. It is nevertheless necessary that a model should take cognizance of the result of extrapolation of the observed region of the V, θ, T characteristics. Phenomenologically, this required a model in which the generation rate initially increases with an increase in field and/or temperature. At a critical field F_c, corresponding to the voltage V_c, the temperature should cease to have appreciable effect on θ, and a similar argument applies to the effect of the field at T_c. It should be noted that the measurements have not been extended up to T_c and V_c as the current became unduly high. It is probable that the straight lines of Figures 1 and 2 should be continued, beyond the experimental points, as curves that approach θ_c tangentially.

The following model would appear to meet these requirements, at least qualitatively. Generation centres are assumed to be present, existing in successive stages of ionization and lying at levels E_1, E_2, E_3 in the band gap. The proportion of the total number of centres N existing in dynamic equilibrium in the depleted space-charge region, at a stage of ionization n_1, is a function of E_1, E_2, E_3, F, and T.

The generation rate will be a minimum when all are at the ionization stage n_1, the assumed state at low temperature and fields.

These centres at E_2, lying nearest to E_1, will provide the maximum contribution to the generation rate. As the temperature or field is increased from the lowest levels the effect will be to increase the proportion at the higher ionization levels. Thus at first the proportion of the centres lying at E_2 will be increased. However, a stage will be approached, as the field or temperature is further increased, at which an appreciable proportion of the centres will be further ionized to level E_3. At this stage

TABLE 1

Constants for Representative Soft Junctions

Diodes	1	2	3
E_{as} in Electron-Volts from 20 to 70 Degrees Centigrade	0·60	—	0·6₂
E_{as} in Electron-Volts from 70 to 205 Degrees Centigrade	0·8	—	—
E_a in Electron-Volts	0·9₃	0·8₅	0·9₅±0·05
$V_c^{0\cdot3}$	4·7	4·3	6·1
Critical Temperature in Degrees Kelvin	590	585	625
B	0·2	0·2	0·16
C	4·0	4·1	3·0
θ_c in Arbitrary Units	2×10^4	9×10^3	8×10^3
θ in Microamperes, 120 Volts, 20 Degrees Centigrade	40	100	0·3

The constants refer to the general equations

$$\theta - \theta_s = A \exp\left\{-\left(\frac{E_a - BV^{0\cdot3}}{kT}\right) - CV^{0\cdot3}\right\}$$

and

$$\theta - \theta_s = \theta_c \exp\left\{-\frac{E_a}{k}\left(1-\frac{V^{0\cdot3}}{V_c^{0\cdot3}}\right)\left(\frac{1}{T}-\frac{1}{T_c}\right)\right\}.$$

E_{as} was obtained from the slope $\partial \ln \theta_s/\partial(1/T)$ in the saturation region, E_a and B from $k\partial \ln (\theta-\theta_s/\partial(1/T) = -E_a+BV^{0\cdot3}$ (Figure 2), C from $\partial \ln (\theta-\theta_s)/\partial V^{0\cdot3} = B/kT-C$ (Figure 1), and T_c and V_c are read directly from Figures 1 and 2. Diodes 1 and 2 are for gallium diffusion into polished slices.

Diode 3 for comparison is a partially B_2O_3 gettered slice.

the normal increase in generation rate consequent upon the increase in temperature will be counterbalanced by the transfer of centres from the more-effective level E_2 to the less-effective level E_3.

The mathematical formulation of this model leads to expressions that can only be solved graphically, even for the grossly simplifying assumption of a uniform field in the space-charge region. The value of the model subsists in the indication it provides that centres with multiple ionization levels are responsible for the excess current in soft diodes.

3. Experimental

All the results reported here refer to *p-n* junctions obtained by diffusing gallium into 5-to-25-ohm-centimetre *n*-type silicon. The silicon consisted of both zone-grown and crucible-grown material from a variety of sources. The silicon slices were pre-oxidized in moist oxygen (25 millimetres of water) to produce a coherent silica film 0·4 micron thick. The slices were individually laid on a flat quartz boat and inserted rapidly into the hot zone of the furnace.

The diffusion was carried out in an open-tube quartz furnace with the source (specially pure Ga_2O_3) at 925 degrees centigrade and the slices at 1200 or 1245 degrees centigrade. The slices were withdrawn from the diffusion furnace at about 20 degrees centigrade per minute. Hydrogen plus water vapour (10-millimetres partial pressure) was used as a carrier gas.

Commercial gases, piped in copper, were dried by passing through activated alumina and humidified by bubbling through deionized water. An initial series of experiments in which the oxygen was passed through a silica-gel trap at −80 degrees centigrade, the hydrogen through activated charcoal at the same temperature, and both gases through porous metal-dust filters, gave precisely the same results as with the untreated gases. In the majority of the experiments therefore untreated gases were used.

The surface concentration of gallium was $2{\cdot}2 \pm 0{\cdot}2 \times 10^{18}$ atoms per cubic centimetre producing junctions 36 microns deep on both sides of the slice.

In the gettering experiments the boric oxide was applied to the surface of the gallium-diffused slice by vacuum evaporation of anhydrous B_2O_3 glass from a heated platinum strip. Uniformity, absence of dust, and thickness were investigated by visual inspection of the interference colour of the film. The slice was transferred rapidly from the evaporation plant to the hot furnace to avoid moisture pick-up. Raymond [1] has found that boric oxide applied by conventional paint-on techniques produces results consistent with those reported here.

Specimens were prepared for measurement by removing the oxide film from the slice in hydrofluoric acid and depositing nickel on both sides from an electrodeless nickel bath. In some runs the nickel was alloyed with the surface by heating for a short period (2 to 3 minutes) at 600 degrees centigrade in pure hydrogen. In a control run in which the nickel contact was not alloyed, there was no significant difference in the reverse characteristics of the diodes so formed. A pattern of mesas, each 2 millimetres in diameter, was made by masking with vacuum wax and etching in CP8. The wax was removed by trichlorethylene vapour, and the exposed junction surface was stabilized by heating in concentrated sulphuric acid for about 2 minutes at 200 degrees centigrade, washing in recirculated deionized water, and baking in air at 240 degrees centigrade for a minimum of 15 minutes. Measurements were made, unless otherwise stated, in dry nitrogen in the dark.

4. Chemical Gettering

Slices of gallium-diffused silicon, chemically or mechanically polished prior to diffusion,

were coated with boric oxide on one or both sides.

The slices were heated in an open-tube quartz furnace in a stream of oxygen. The effect of time, temperature, and moisture content of the oxygen, is shown in Figure 3 and Table 2. The data, with one exception, refer to measurements on mesas formed on the coated side of the slices. The breakdown voltage was arbitrarily specified as the voltage at which the reverse current of a 2-millimetre-diameter mesa reached 10 microamperes. In all cases the breakdown voltage of the junction prior to treatment was considerably less than 100 volts.

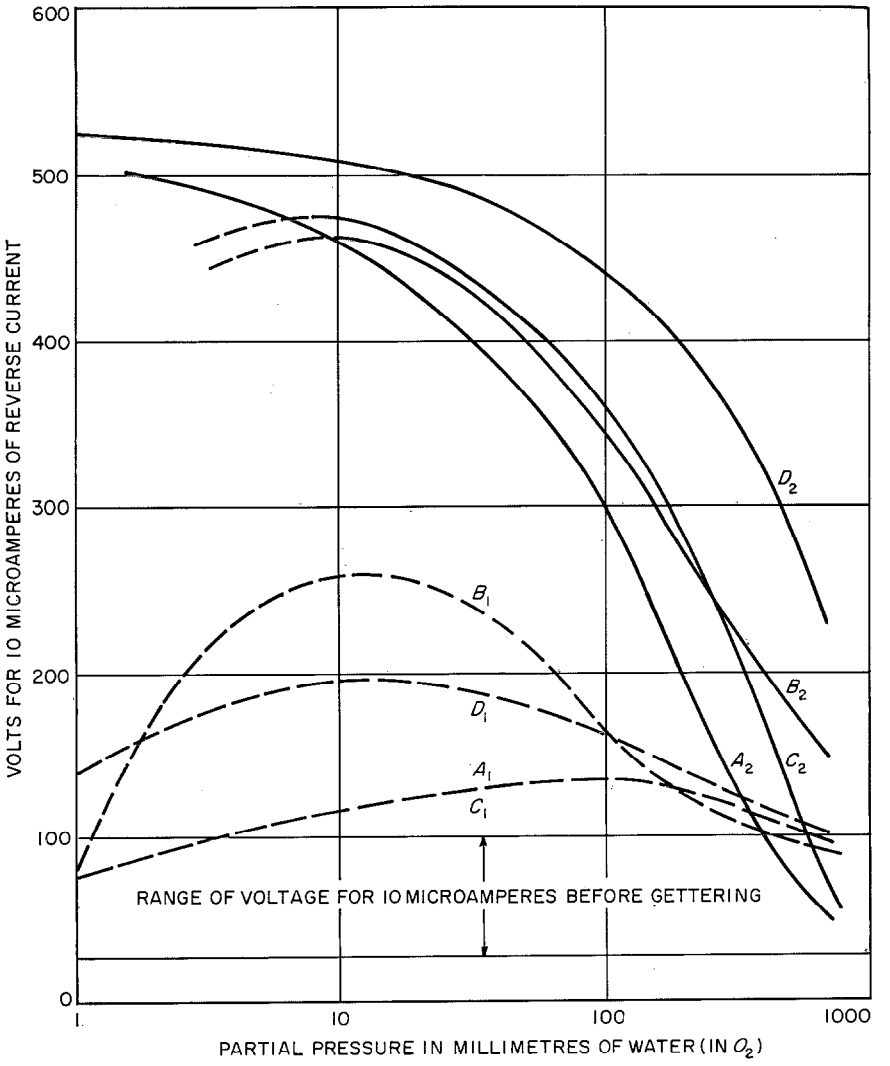


Figure 3—Junction quality as a function of humidity during boric-oxide gettering. Slices with soft junctions 35 microns deep on both sides, were coated on one side with boric oxide and gettered at 1200 degrees centigrade in moist oxygen. The letters on the curves indicate the intervals at which the boric-oxide glass was replaced: *A*, no replacement; *B*, after 3 hours; *C*, after 2 hours; and *D*, after 2 and 3 hours. The total gettering time was 4 hours in all cases. Subscript 2 indicates junctions on the boric-oxide side and 1 signifies junctions on the opposite face.

TABLE 2

QUALITY OF JUNCTION FOLLOWING BORIC-OXIDE GETTERING

Boric-Oxide Gettering		Voltage Range of All Tests	Proportion of Tests in Which Voltage for 10 Microamperes of Reverse Current Exceeded Specified Value	
Temperature in Degrees Centigrade	Time in Hours		Proportion	Volts
1200	1/4	100–400	5/8	300
1200	1/2	310–380	3/4	310
1200	3/4	324–420	15/15	324
1200	1	384–444	5/5	384
1200	1-1/2	456–540	5/5	456
1090	1/4	170–260	7/9	180
1090	1-1/2	450–480	8/8	480

For reference, gallium-diffused ungettered junctions were tested between 40 and 80 volts and none operated at 100 volts with the required 10 microamperes of reverse current.

The quality of the junction on the side of the slice opposite to that on which the boric oxide had been deposited varied from run to run. This was found to be so for slices ranging from 180 to 400 microns in thickness. Traces of contamination left on, or subsequently entering, the uncoated face were found to be responsible for this variability. It was shown that the impurities were deposited on the slice during dissolution of the silica film in hydrofluoric acid, and that further contamination during the subsequent heat treatment was virtually unavoidable. Coating the untreated face of the slice with thermally grown oxide provided no protection.

An experiment in which contamination was deliberately applied to the oxidized surface of a hard junction, which was subsequently reheated to 1200 degrees centigrade for a short period, showed that the oxide film was readily permeable to the likely contaminants.

Slices had both faces polished prior to the gallium diffusion. Boric oxide was applied on both sides of the slice, which was inserted straight into the hot zone of the furnace and then withdrawn rapidly. The atmosphere was of O_2 at a pressure of 25 millimetres of water. Measurements were made on 2-millimetre-diameter mesas.

4.1 MECHANISM OF CHEMICAL GETTERING

The results are consistent with the view of Goetzberger and Shockley [12] that a liquid phase of acidic oxide combines with the electro-positive impurity atoms. The surface concentration of the impurity remains very small until the boric oxide is excessively diluted with silica. It is unlikely that the formation of a solid film of silica is responsible for the cessation of gettering that occurs more rapidly the greater the water content of the oxygen. It has already been demonstrated that silica is readily penetrated by these impurities, and it is more likely that displacement of the equilibrium

$$M + O_2 \rightleftharpoons M \cdot \text{oxide} + B_2O_3 \rightleftharpoons M \text{ oxide} \cdot B_2O_3$$

towards the left, by the increased competition for the B_2O_3 by the silica is responsible.

The effectiveness of short times and low temperatures (Table 2) confirms the view that fast-diffusing impurities are being removed. Amongst such elements iron, copper, gold, and silver are known to have appropriate diffusion constants.

The difficulty of gettering a junction on the remote face of a slice, with boric oxide on the opposite face, was not a function of the thickness of the slice. It was evident that the rate of transport through the slice was adequate, as in those runs in which attempts to preserve the cleanliness of the remote face were successful, the quality of both junctions was excellent. However, the difficulty in completely removing any source of contamination made it clearly desirable to coat both faces of a specimen with the gettering agent.

A less-exhaustive series of experiments showed that a film of phosphorus glass behaved in a qualitatively similar fashion as a gettering agent. The glass was predeposited by passing a stream of dry oxygen carrier gas saturated

by phosphorus pentoxide maintained at 230 to 270 degrees centigrade over the oxide-free silicon slice at 1050 degrees centigrade. The subsequent heat-treatment was essentially the same as that used during boric-oxide gettering.

5. Physical Gettering

Chemical gettering provides an extremely effective method of obtaining hard junctions, but it is not always conveniently applicable to device production. Where the device can be so designed that the boric oxide or phosphorus pentoxide simultaneously produces the appropriate junctions, and acts as a getter, the chemical process is eminently suitable. But concomitant boron or phosphorus doping is unavoidable, and the whole surface of the slice must be coated if reliable gettering is to be obtained.

The new process of physical gettering described below is felt to be capable of more-general application.

The soft junctions discussed above were the result of gallium diffusion into mechanically or chemically polished slices. The mechanical polishing programme was designed to produce a surface with a damaged layer less than 1 micron thick, and the chemical polishing to remove completely the damaged surface layer.

A series of slices was so treated as to produce a surface layer of material of considerable thickness with a high density of lattice damage. The slices, gallium diffused after suitable cleaning treatments, had reverse characteristics similar to those of the chemically gettered slices.

Soft junctions that had received the same surface treatment (after diffusion) and were subsequently reheated for a time short compared with the diffusion time, showed improvements of the same order. Table 3 shows typical results for a single batch of slices which, after a standard gallium diffusion, were divided into 3 groups. Batch *A* was measured without further treatment, batch *B* was reheated without further treatment, and batch *C* was given a physical gettering treatment. Junction quality is defined by the voltage at which the reverse current of a 2-millimetre mesa reaches 10 microamperes.

TABLE 3

EFFECT ON SOFT JUNCTIONS OF SURFACE DAMAGE AND REHEATING

Batch	Slice Preparation	Junction Quality
A	Gallium Arsenide Diffused	30 Volts
B	Gallium Arsenide Diffused and reheated	16 Volts
C	Surface Prepared for Physical Gettering and Reheated (as B)	480 to 540 Volts

Typical results of similar treatments are shown in Figure 4.

5.1 MECHANISM OF PHYSICAL GETTERING

The similarity between the requisite condition for, and the results of, chemical and physical

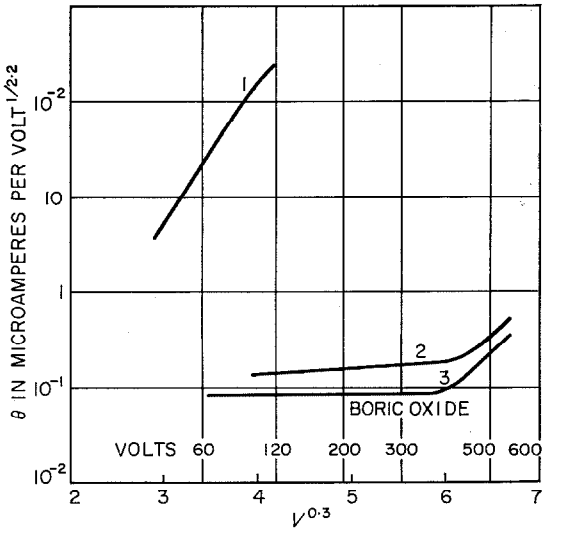


Figure 4—Effect on junction quality of surface preparation prior to gallium diffusion. Curve 1 is for a mechanically polished slice. Curve 2 is physically gettered and curve 3 is boric-oxide gettered. The gallium diffusion was 35 microns deep in all cases.

gettering, suggests that the mechanisms are themselves similar, involving in both cases the transfer of impurities from the bulk, to the vicinity of the surface where they become bound.

Kikuchi and Itzima [16] observed that grinding the surface of a germanium sample and re-heating for 1·5 hours at 650 degrees centigrade produced a substantial increase in the lifetime. They suggested that precipitation of copper at lattice defects was occurring. Pugh and Samuels [17] have shown that the primary result of abrading a germanium surface is the introduction of an extensive system of micro-cracks extending below the broken surface layer.

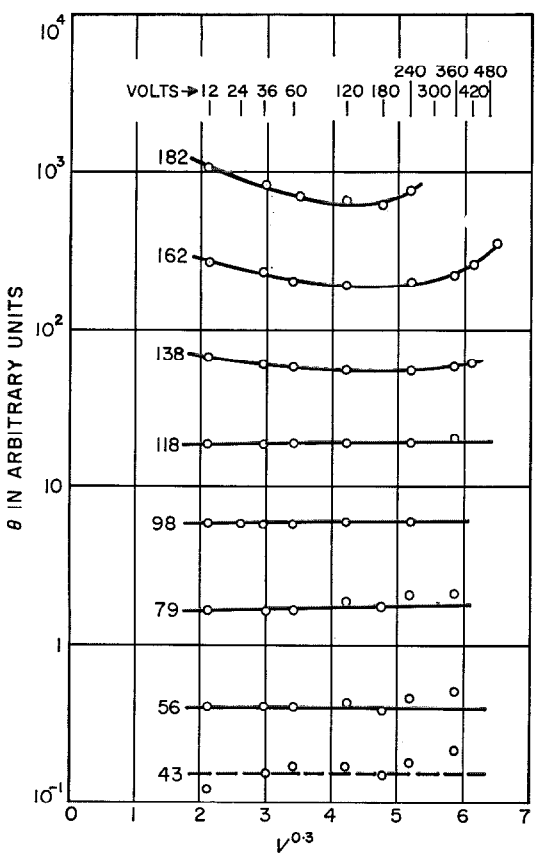


Figure 5—Reverse characteristics of typical gettered junction. The temperature in degrees centigrade is noted for each curve.

The following picture emerges: specimens that have been subjected to a long-term temperature treatment are rich in a variety of impurities. These impurities may be present in saturated or supersaturated solution, and as precipitates. From experiments in which soft diodes with undamaged surfaces were cooled from elevated temperatures rapidly or slowly, without significant change in their diode characteristics, it appears that further precipitation of the impurity cannot be induced in the absence of surface damage.

Thus, in soft diodes with undamaged surfaces, the impurities are distributed at room temperature in a manner that cannot readily be disturbed by further heat treatment. They may be present as precipitates that are dissolved and reprecipitated on heating and cooling, or they remain in solution, saturated at high temperature, supersaturated at room temperature.

On the other hand, a new process can occur if the surface of the specimen is damaged before reheating. Fuller and Ditzenberger [19] and others, have shown that the diffusion rate of many transition elements in germanium and silicon is very dependent on the dislocation density. In the presence of dislocations, the vacancy density permits equilibrium to be maintained between substitutional and interstitial atoms. Whilst interstitial solubility is normally an order of magnitude lower than substitutional, rapid diffusion ($D_{1000} = 10^{-5}$ centimetre^2 second^{-1}) can occur only interstitially.

During the cooling of a specimen with a damaged surface, precipitation will occur most readily at the broken surface, and around the micro-cracks. A concentration gradient is set up, in which the impurity will diffuse from the bulk to the surface until the flow is quenched by the drop in diffusion constant as the temperature falls. In view of the low activation energy of the interstitial diffusion process, flow will persist during the greater part of the cycle. Chemical gettering may also be partly

dependent on the defects present in silicon surfaces containing a high boron concentration. Only under such conditions can substitutional-interstitial equilibrium be maintained, permitting rapid interstitial diffusion to occur.

This view of the physical process is in keeping with an experiment in which two samples of silicon with soft gallium junctions 10 and 35 microns deep respectively have been treated to produce shallow surface damage.

After subsequent heat treatment, the deep junction had become hard while the shallow one was further degraded. Clearly the shallow junction was within the layer affected by surface damage. After reheating, further changes had occurred that were presumed to involve the removal of impurities from the region of the deep junction and precipitation in the vicinity of the shallow one.

Thomas [20] has shown that the mode of precipitation of copper in silicon depends on the surface treatment. Samples saturated at 1100 degrees centigrade with copper and cooled to room temperature in 10 seconds exhibit characteristic etch figures. The number and depth of the precipitates responsible for these etch figures depends on the surface preparation of the silicon. The thickness of the layer, in which precipitation at a density of 10^5 to 10^6 centimetre^{-2} occurred, was of the same order as that of the damaged layer. If the silicon surface had been previously chemically polished, the precipitation occurred uniformly throughout the material as a density of approximately 10^3 centimetre^{-2}. Thermal oxidation or prolonged annealing at elevated temperatures removed the worked layer present after mechanical polishing.

6. Nature of Hard Junctions

The characteristics of junctions produced by physical and chemical gettering are very similar. Over most of the range of reverse bias, θ_s is constant (Figure 5). The activation energy of the generation process is very close to $E_g/2$ below 70 degrees centigrade, and increases to about 0·75 electron-volt at higher temperatures (Figure 6). At high temperatures the initial decrease in θ_s with increasing voltage, is the result of a significant contribution by the voltage-independent diffusion current. The identification of a second process with an activation energy of 0·75 electron-volt must therefore be treated with reserve, as the diffusion current contribution with an activation energy of 1·2 electron-volts cannot be separately estimated.

The junctions have been designed to break down (avalanche mode) at about 600 volts. It is therefore not possible to separate avalanche effects in the high-voltage region from voltage-sensitive generation of the type found

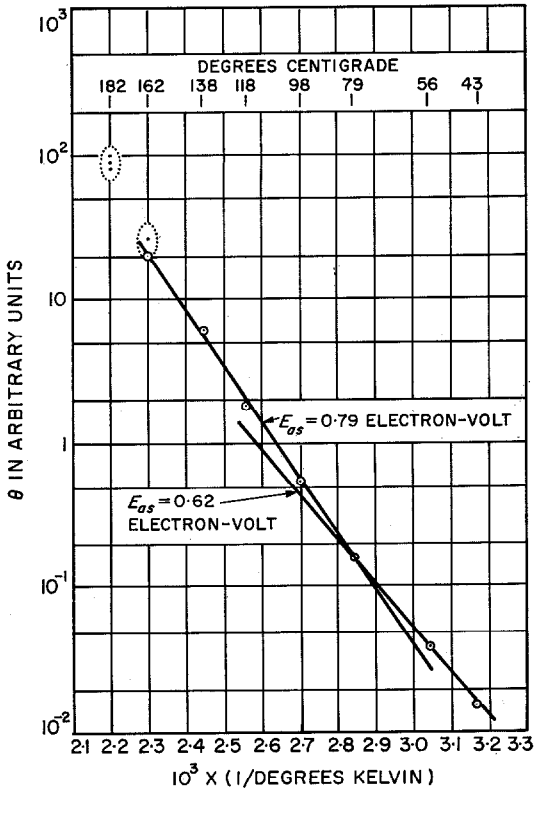


Figure 6—Reverse characteristics of the hard junction of Figure 4.

in soft junctions. It is clear however that any such effect had been largely eliminated by gettering.

The difference between the surfaces of hard and soft diodes is of interest. It is difficult to produce quantitative data relating to surface current, as diodes with bad surfaces are notoriously difficult to reproduce or even to measure. Comments must therefore be based on general experience. It has been noted that the harder the diode the smaller the absolute increase in the surface current as the surface conditions are degraded.

The reverse characteristics of hard junctions are therefore characterized by a current proportional to $V^{0.46}$, a temperature coefficient given by $(\partial \ln \theta_s/\partial(1/T) = 0{\cdot}6/k$ in the range from 20 to 70 degrees centigrade, and a relative insensitivity to surface contamination.

7. Conclusion

Crucible and float-zone silicon from a variety of sources has been investigated. No significant differences have been observed in the properties of well-gettered diodes made from these samples. Silicon treated in this way therefore provided a conveniently uniform starting material for other investigations.

The profound effect of apparently minor alterations in surface preparation, diffusion procedure, and cooling rate on lifetime and reverse junction current may well explain the diversity of the results reported by various investigators. The effect of heat treatment on recombination centre density is determined by the nature of the sample surface. Unresolved differences between results of such experiments can be explained by the differing amounts of surface damage initially present.

The extent to which chemical gettering occurs during the production of boron- and phosphorous-diffused structures will depend on many factors. Failure to take this into account in the design of the diffusion schedule can lead to a fluctuating yield of hard devices.

Further work is necessary to identify the elements responsible for the softness of some diodes and to discover the stage at which they enter the silicon. As long as the conventional methods of material preparation and diffusion are used it will probably be impossible to eliminate completely contamination of the silicon. Methods for removal of the contaminants are therefore essential.

It is fortunate that the two methods described herein are available. In the case of shallow-junction devices it may be necessary to adopt the chemical method, when a damaged surface layer cannot be tolerated. It is important to note, however, that whilst the impurities are removed from the system with the boric-oxide-silica glass, they remain in the surface after physical gettering. Long-term changes in lifetime, even at room temperature, reported for example by Gallagher [15] may be associated with release of the impurities from the surface layer. Precautions can be taken to avoid this contingency by removing the surface layer after physical gettering.

The additional restraint imposed on the designer if efficient gettering is to be included as an integral part of the production process, may demand some modification of this process. The gain in device quality and reproducibility would appear to warrant such a change.

8. Acknowledgments

The authors wish to express their gratitude to Miss B. Williams and Mr. R. Amos who have undertaken much of the experimental work, and to Mr. D. J. Dane-Thomas for many helpful discussions during the course of this work.

9. References

1. F. J. Raymond (Private Communication).

2. Maurice Bernard, "Mesures en Fonction de la Température du Courant dans les Jonctions de Germanium n-p," *Journal of Electronics*,

volume 2, First Series, pages 579–597; May 1957.

3. E. L. Nolle and G. L. Galkin, *Fizika Tverdogo Tela*, volume 3, number 8, page 2350; 1961.

4. E. M. Pell, "Reverse Current and Carrier Lifetime as Function of Temperature in Germanium Junction Diodes," *Journal of Applied Physics*, volume 26, pages 658–665; June 1955.

5. E. M. Pell and G. M. Roc, "Reverse Current and Carrier Lifetime as Function of Temperature in Silicon Junction Diodes," *Journal of Applied Physics*, volume 27, pages 768–772; July 1956.

6. A. R. Plummer, "The Effect of Heat Treatment on the Breakdown Characteristics of Silicon pn Junctions," *Journal of Electronics and Control*, volume 5, First Series, pages 405–416; November 1958.

7. D. J. Sandiford, "Heat Treatment Centers and Bulk Currents in Silicon p-n Junctions," *Journal of Applied Physics*, volume 30, pages 1981–1986; December 1959.

8. C. T. Sah, R. N. Noyce, and W. Shockley, "Carrier Generation and Recombination in P-N Junctions," *Proceedings of the IRE*, volume 45, pages 1228–1243; September 1957.

9. L. A. Belova, A. H. Kovalev, and L. A. Penin, *Fizika Tverdogo Tela*, volume 2, number 10, page 2647; 1960.

10. A. V. Chevychelov, *Fizika Tverdogo Tela*, volume 1, number 8, page 1205; 1959.

11. A. G. Chynoweth, "Ionization Rates for Electrons and Holes in Silicon," *Physical Review*, volume 109, pages 1537–1540; March 1, 1958.

12. A. Goetzberger and W. Shockley, "Metal Precipitates in Silicon p-n Junctions," *Journal of Applied Physics*, volume 31, pages 1821–1824; October 1960.

13. V. J. Fistul and O. B. Orzhevskii, *Fizika Tverdogo Tela*, volume 2, number 9, pages 221–224; 1960.

14. Ibid. *Fizika Tverdogo Tela*, volume 3, number 4, page 1158; 1961.

15. C. J. Gallagher, "Thermal Defects in Silicon," *Physical Review*, volume 100, Abstract, page 1359; 15 November 1955.

16. M. Kikuchi and S. Itzima, *Journal of the Physical Society of Japan*, volume 12, page 824; 1957.

17. E. M. Pugh and L. E. Samuels, *Journal of the Electrochemical Society*, volume 108, number 11, page 1043; 1961.

18. Ibid. *Journal of the Electrochemical Society*, volume 109, number 5, page 409; 1962.

19. C. S. Fuller and J. A. Ditzenberger, 'Effect of Structural Defects in Germanium on Diffusion and Acceptor Behavior of Copper," *Journal of Applied Physics*, volume 28, pages 40–48; January 1957.

20. D. J. Dane-Thomas (Private Communication).

Cyril Drake was born in Barking, Essex, England, in 1921 and received the London University degree of B.Sc. in chemistry in 1940.

During World War II, he worked with the British Ministry of Supply on explosives and anti-fouling paints.

In 1946, he joined Standard Telecommunication Laboratories and has worked on magnetic materials, X-ray crystallography, selenium rectifiers, printed circuits, semiconductor devices, and solid-state circuits.

Mr. Drake is a member of the Faraday Society.

Kenneth L. Ellington was born in 1923 in Wanne-Eickel, Germany. In 1951, he received the London University degree of B.Sc. in physics.

He joined Standard Telecommunication Laboratories in 1951, working on X-ray crystallography and point-contact switching devices, being engaged in problems of breakdown and surface protection in semiconductor devices.

Defects in Vapour-Grown Silicon

K. O. BATSFORD
D. J. D. THOMAS
Standard Telecommunication Laboratories Limited; Harlow, Essex, England

1. Introduction

A very-important development in device technology is the deposition from a vapour of materials, chiefly silicon, in single-crystal form. The substrate, usually a single crystal of the depositing material, is held at a temperature that may be several hundred degrees below the melting point of the deposit. Diffusion of impurities is considerably reduced at these low temperatures so that, by suitably doping the vapour prior to deposition, *p-n* junctions with predetermined impurity profiles may be grown.

Crystal imperfections play an important role in the mechanical and electrical characteristics of semiconductor materials and devices. In particular, dislocations may produce electrical effects in themselves and perhaps more importantly act as nuclei for the precipitation of foreign atoms that may be present in the parent matrix [1]. Such precipitates may have considerable effects on the properties of *p-n* junctions. Much information has been accumulated on the influence of growth conditions on the number and type of imperfections found in semiconductor crystals grown from the melt. In this paper an investigation into defects in vapour-grown silicon layers is reported. At the temperatures used for growth, the mobilities of silicon and impurity atoms, and of defects are much lower than at the melting point and this may well have an effect on both the type of the defect and its mode of formation. Several methods have been employed to investigate the defects including metallography, X-ray and electron diffraction, and transmission electron microscopy. The defects found in the layers have been studied and methods developed for controlling their numbers.

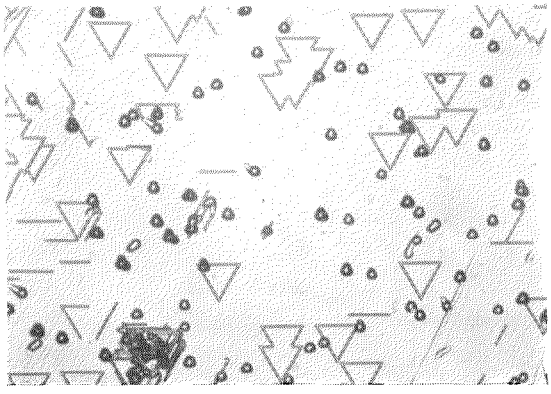


Figure 1—Etched surface of (111) epitaxial layer. (× 250)

2. Preparation of Samples

Most of the work has been carried out on layers grown on mechanically polished substrates. The samples were initially lapped with silicon-carbide abrasives and subsequently polished with alumina powders, the final particle size being 1 micron. The slices were then cleaned ultrasonically in mixtures of hydrofluoric acid and methanol after vapour degreasing.

The layers were grown at rates between 0·5 and 1 micron per minute at 1200 degrees centigrade by the decomposition of either silicon tetrachloride or silane. The hydrogen used as ambient was purified by diffusion through a palladium-silver alloy. The water and oxygen concentrations were below the detection limit of 0·1 part per million.

3. Examination of Layers

3.1 Line Defects

Most samples of epitaxial layers reveal, on etching in either Dash or iodine etch, line etch figures that in (111) surfaces often assume the form of equilateral triangles as may be seen in Figure 1. Normal dislocation etch pits can also be seen. The triangles on a particular sample are all the same size and their sides are parallel to ⟨110⟩ directions. The length of a side of the triangle was found to be about 1·2 times the layer thickness [2]. On (110) layers

the lines meet at angles of 70 degrees, on (100) layers the angles are 90 degrees and on (112) layers isosceles triangles with angles of 44 and 68 degrees are seen. On repeated etching of (111) surfaces, the centre of the triangle becomes progressively smaller until, when the depth of etching is the same as the layer thickness, the triangular shape become a point. On further etching the figures gradually disappear. See Figure 2.

From these facts, it is concluded that the etch figures are due to 2-dimensional defects that lie parallel to (111) planes and are introduced when growth commences. This is very well shown in Figure 3, where a sample has been cleaved to reveal the $(11\bar{1})$ surface and subsequently etched. The apex of the tetrahedron has not been exactly intersected, and the interface between substrate and deposit is well delineated as a p-n junction.

To investigate the origin of these defects, some annealing experiments were carried out. The samples were heated for various times and temperatures in pure hydrogen before growth was commenced at 1200 degrees centigrade.

As shown in Table 1, at the various growth temperatures, the defect density is seen to fall off, and under our normal growth conditions a 15-minute pre-anneal reduces the defect density below 100 per square centimetre.

TABLE 1
EFFECT OF PRE-ANNEALING ON DEFECT DENSITY

Annealing Time in Minutes	Number of Defects per Square Centimetre			
	Growth Temperature in Degrees Centigrade			
	1210	1150	1050	930
1	6×10^4	$2{\cdot}5\times10^5$		
2	8×10^3	8×10^4	2×10^5	
3	2×10^3			
4		5×10^3	3×10^4	
6			1×10^4	
10				7×10^5
15	1×10^3			1×10^5
30				9×10^4
60				8×10^4

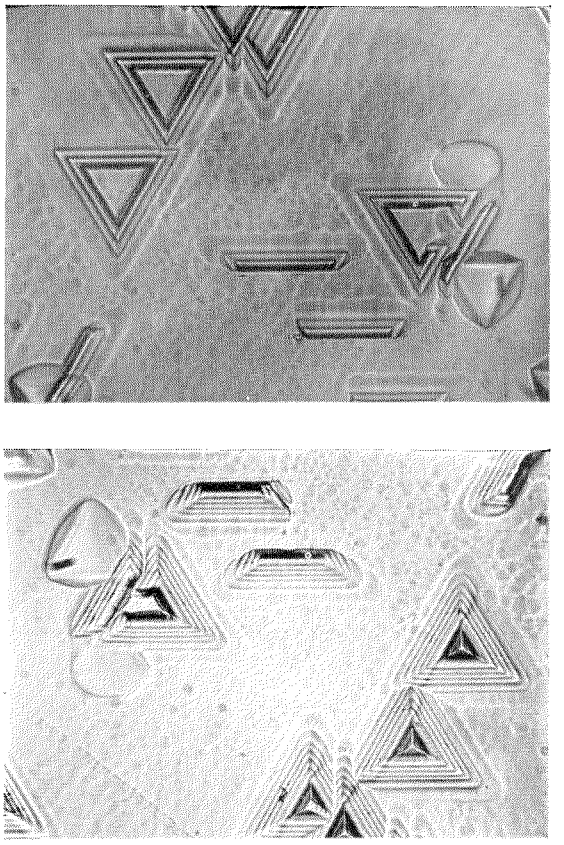


Figure 2—Effect of repeated etching on triangular structure. (× 1000)

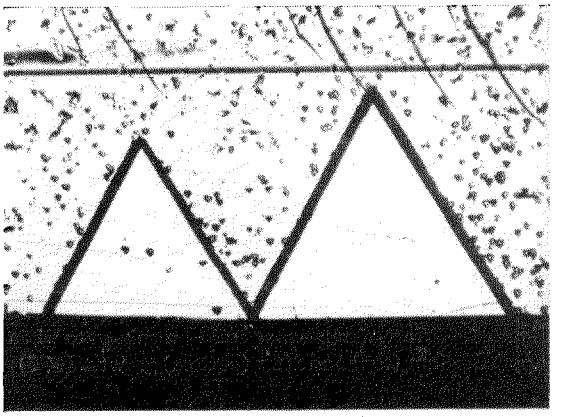


Figure 3—Etched $(11\bar{1})$ surface of epitaxial layer. (× 180)

These results suggest that the defects are caused either by an imperfection in or an impurity on the surface. Annealed samples that were covered with a thin (20-Ångström-unit) oxidized layer and also chemically polished samples, had about 100 defects per square centimetre when grown in conditions that would normally introduce about 10^5 per square centimetre. The activation energy of annealing for the defects, about 35 kilocalories per mole, is much lower than has been found for oxide removal in hydrogen (about 100 kilocalories per mole) [3]. This suggests that the cause of the line defects is physical rather than chemical in mechanically polished layers.

It was observed, however, that when chemically polished substrates were deliberately contaminated, such as by impure water or acetone, so that a visible film was deposited on the substrate, the grown layer showed large numbers of defects. Figure 4 shows the effect of a tap-water stain on a substrate before and after growth, and in Figure 5 the etched surface of 20-micron-thick silicon deposits at the edges of water and acetone stains respectively can be seen. By allowing acetone vapour to condense and re-evaporate from a substrate prior to growth, large numbers of line etch figures are generated, while the more-coarse water deposit gives rise to twinning, dislocations, and line figures.

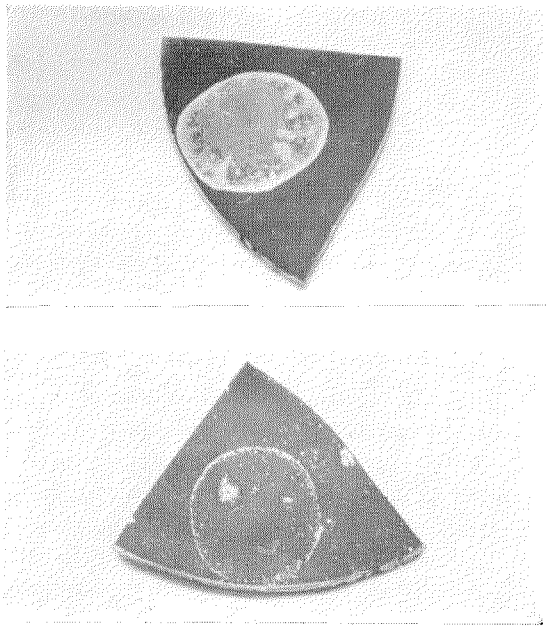


Figure 4—Effect of tap-water stain on substrate when applied before (top) and after (bottom) growth. (× 3)

3.2 Dislocations

It is immediately noticeable that the number of dislocation etch pits in layers grown on clean well-annealed substrates is substantially the same as in the substrate itself. It should be noted that dislocations lying in the (111) plane parallel to the surface, that is, not intersecting the surface, are not propagated and on etching $(11\bar{1})$ surfaces, fewerdisloc ations are

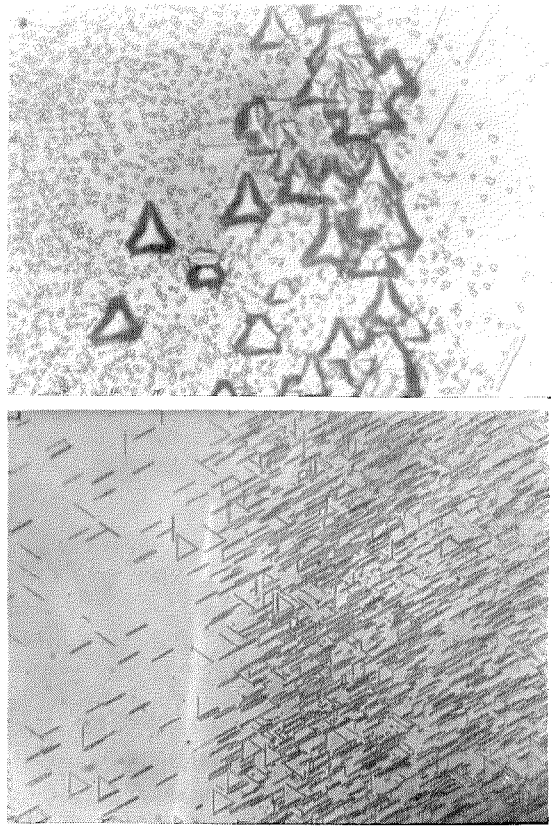


Figure 5—Effect on growth of water (× 500) (top) and acetone (× 330) (bottom).

seen in the layer than in the substrate. In contaminated layers the number of dislocations may be very large as was shown in Figure 5. In insufficiently cleaned substrates occasional bunches of dislocations are observed, which are also probably due to contamination.

Lineage boundaries are easily shown to propagate, and single dislocations can be shown to propagate by the following experiment. A slice was scratched with a diamond and annealed for 15 minutes at 1200 degrees centigrade before growing a layer on it. During the anneal, stresses in the scratch caused dislocations to move into the silicon along the {111} slip planes. These dislocations were propagated and considerable numbers of line defects were also introduced as is evident in Figure 6. In epitaxial layers that have been highly doped with boron or phosphorus during growth, dislocation nets, produced as a result of lattice mismatch and lying in (111) planes, have been observed. This is similar to the effect found after diffusion of these impurities into bulk silicon [4]. They may be observed either by the Lang transmission X-ray technique [5] or by etching as in Figure 7.

3.3 Polycrystallinity and Twinning

Polycrystallinity and twinning can be conveniently shown up by X-ray back reflection, and Figure 8 shows examples of random polycrystallinity, texture, microtwinning, and a twin lamella. These can be produced by unclean growth conditions or by growth at too low a temperature. Polycrystalline regions expose planes other than (111) and these, growing faster, appear as hillocks. Hillocks are also produced if the substrate is etched rather than polished chemically or mechanically, prior to deposition. From the patterns obtained, these imperfections are found to arise at the sites of etch pits, as shown in Figure 9. From etch-pitted regions X-ray photographs similar to Figure 8*C* are obtained. It is also noted that in the immediate etch-pit region, extra concentrations of triangles are found, and this gives a pointer to the reason for the formation of these structures. In thin layers that are smaller than 20 microns, no extra diffraction effects are obtained from regions with triangles, but on thick layers of 300 microns triangular areas of enhanced diffraction within the usual X-ray diffraction spots correlate on a 1:1 basis with the fault structures.

4. Stacking Faults

4.1 Nature of Line Defects

From etching studies, by X-ray, and by selected-area electron diffraction measurements it was found that the orientation inside

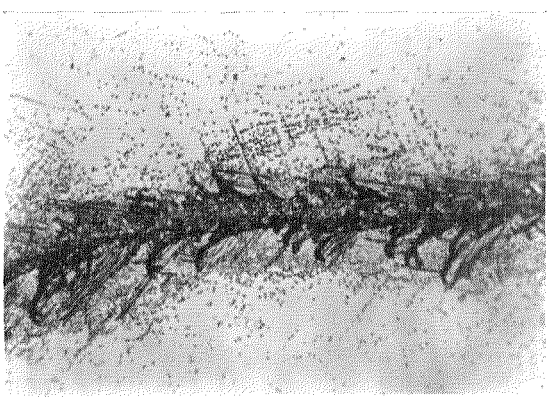


Figure 6—Propagation of dislocations. (× 500)

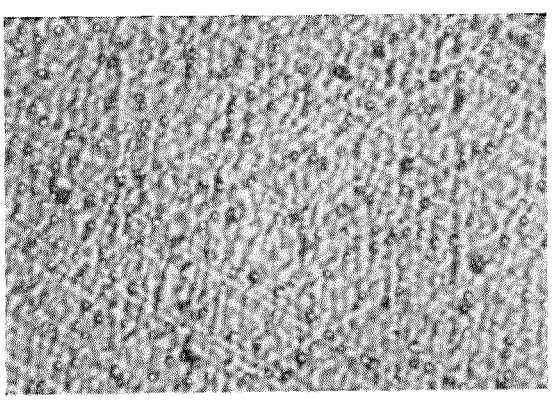


Figure 7—Dislocations in a layer highly doped with boron. (× 1000)

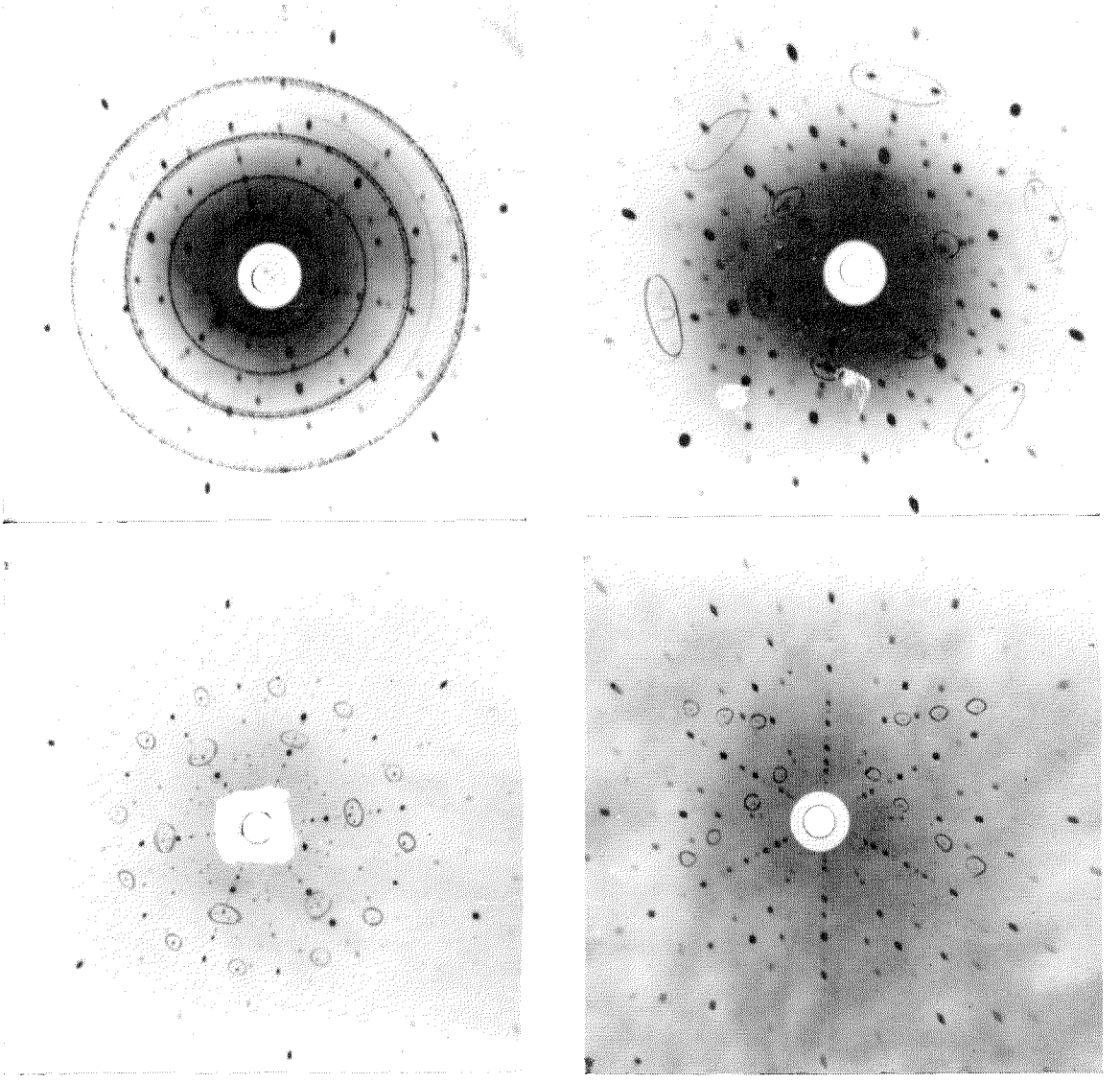


Figure 8—X-ray reflection photographs showing: (top left) polycrystallinity, (top right) texture, (bottom left) microtwinning, and (bottom right) twin lamella.

the triangles on (111) surfaces is the same as that on the outside. This leads to the conclusion that the two-dimensional defects that are parallel to the {111} planes must either be stacking faults or an even number of twins.

To investigate this further some experiments using transmission electron microscopy were carried out [6]. Defects in a crystal lattice produce diffraction contrast and there is a relationship between the Burgers or displacement vector and the Bragg reflection producing the image. Simply no diffraction contrast occurs if the Burgers vector lies in the reflecting plane. In Figure 10 the effect of tilting the specimen is seen. In A, the sides a of the triangles lie in ⟨110⟩ directions; at the ends of the

single faults *b* lines may be seen that lie in ⟨110⟩ directions; at *c* there are some individual lines. In *B* where the specimen has been tilted the diffraction contrast has changed. The contrast due to the faults at *a* has disappeared leaving only a line *b* at the end.

The fringe contrast is shown more clearly in Figure 11. The sample was at such an angle to the beam that faults *a* were not in the diffracting condition. The difference in the width of the various faults is due to changes in thickness of the sample. As in the previous figure there are a number of line defects, some of which are associated with fringes. There is also a line defect at *d* where two faults meet. Selected-area diffraction patterns taken adjacent to and across the faults are exactly similar and show no evidence of extra spots.

From this it is concluded that the defects are stacking faults for the following reasons.

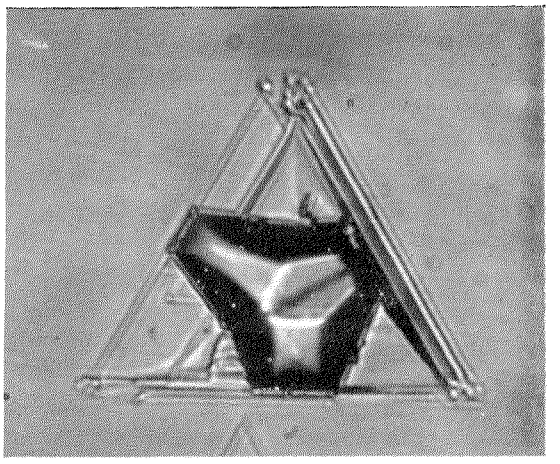


Figure 9—Defects produced at etch pits. (× 1300)

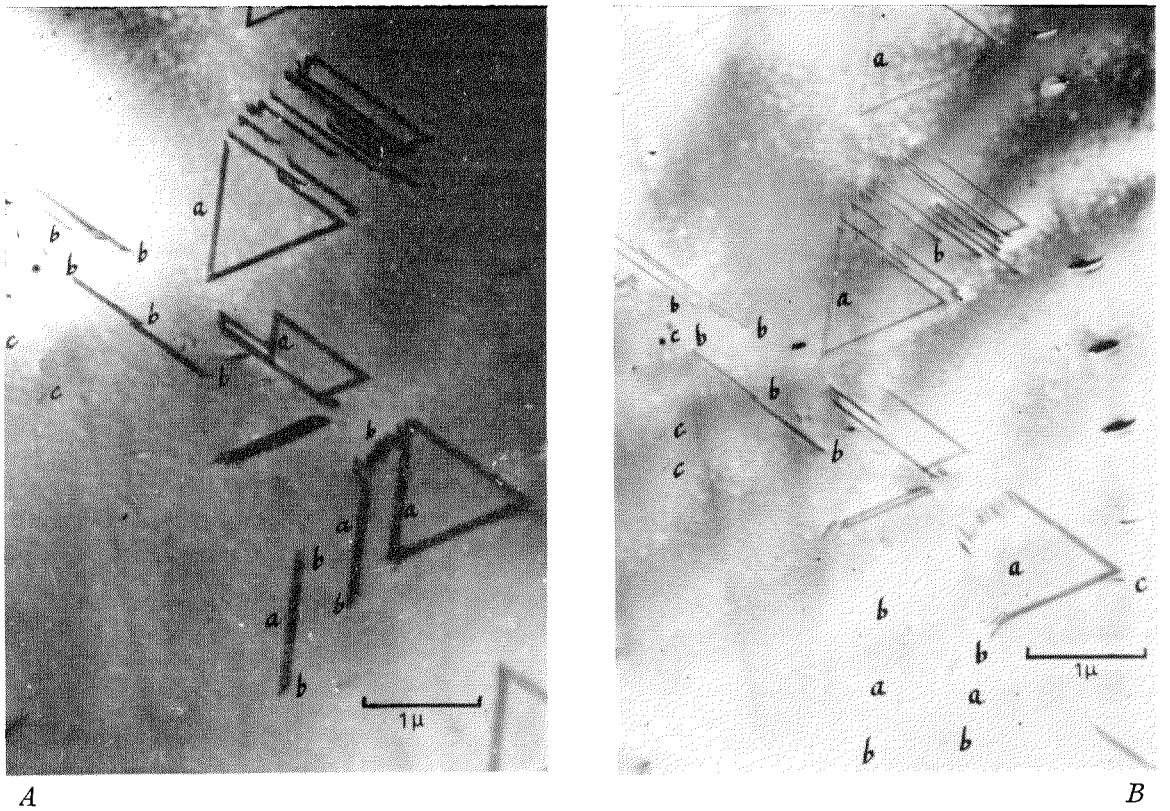
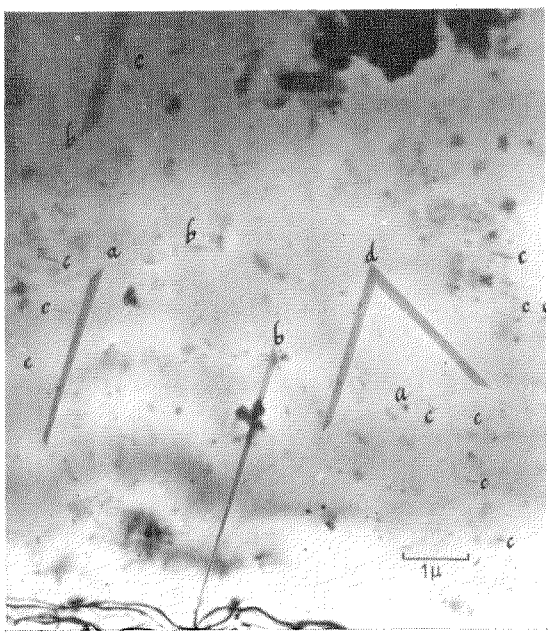


Figure 10—Transmission electron micrograph showing stacking faults and dislocations. (× 17 000)

Figure 11—Transmission electron micrograph showing stacking faults and dislocations. (× 10 000)



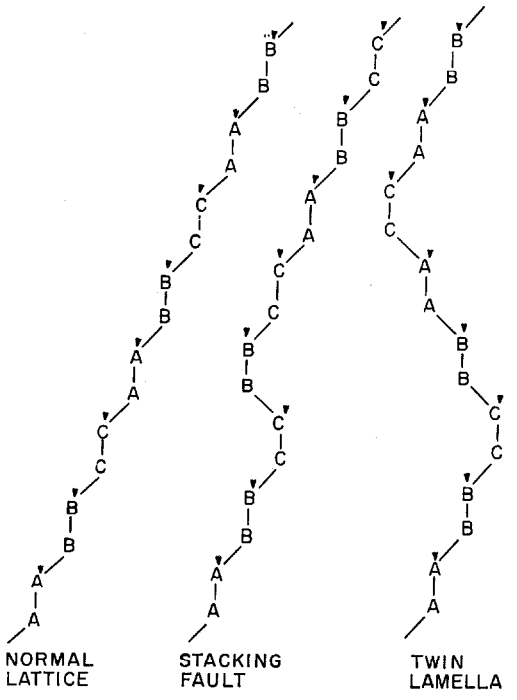


Figure 12—Diagram showing a stacking fault and twin lamella in the diamond structure.

Firstly, stacking faults give rise to diffraction contrast of the type observed. Secondly, single faults end at partial dislocations which are also observed at *b* in Figure 11 and 12. Thirdly, a stair-rod dislocation is produced when two stacking faults on different {111} planes meet and this too has been observed (*d* in Figure 11). Fourthly, multiple twins would produce extra spots on diffraction patterns and these have not been observed. The line contrast at *c* in Figures 10 and 11 is due to individual dislocations.

A stacking fault in the diamond lattice is produced when an extra (111) plane is inserted into or removed from a crystal and is shown schematically in Figure 12. This changes the stacking sequences of the planes from say *abcabcabc* to *abcbcabc*, that is, over the stacking fault region the symmetry is hexagonal. A stacking fault is also a pair of twins on adjacent planes and the boundary has surface energy about twice that of a single twin boundary. A stacking fault is also always associated with partial dislocations. Because of these properties stacking faults are expected to have effects on both dislocations and impurities.

Dislocations may be prevented from moving through the boundary due to the line tension of the latter and they may be anchored by it. The effect on dislocations is clearly shown in Figure 3 where the boundary has prevented dislocations passing into the tetrahedron during growth of the layer. This is analogous to the Dash technique for growing dislocation-free crystals from the melt by having a very thin seed [7]. In the present case the tetrahedron bounded by stacking faults starts from a point and grows outwards producing substantially perfect material. Dislocations may also be pinned by the partial dislocation at the end of a single stacking fault as shown in the transmission electron micrograph of Figure 13.

The stacking fault energy may be lowered by the presence of impurities as is the case in many metal alloy systems. The effect of

stacking faults on copper precipitation may be observed in Figure 14. The copper is preferentially precipitated at the stacking fault and the fault itself is broken up during this process. This type of precipitation in the fault may have a considerable effect on devices made from epitaxial layers.

4.2 Origin of Defects

It has been shown that stacking faults may be introduced into vapour-grown silicon layers and that the great majority are produced at or near the substrate interface. It has also been shown that the number of defects depends on both the physical perfection of the surface and the amount of contamination on the surface. It is suggested that the initial growth commences at specific nuclei and that the origin of a nucleus is either a physical or chemical imperfection. It is probable that these nuclei are slightly misorientated and that, when they meet, a stacking fault is produced. If the misorientation becomes too large other defects may be produced possibly as a result of the strain introduced when the growing nuclei meet. This is particularly the case when etch pits are present for, if the layers grow out from the side of the pit, considerable distortion will be produced when they meet and highly twinned or polycrystalline regions will result. It is interesting to note that polycrystalline regions grow faster than the (111) orientation of the layer as a whole.

Further study of the nucleation process is needed if a more-complete understanding of the mechanism of defect formation is to be obtained.

5. Conclusions

From the reported experiments it is clear that, provided adequate precautions are taken, vapour-grown layers of silicon have defect concentrations similar to the substrate material. To produce these highly perfect layers it is necessary to clean the substrate effectively, to anneal before growth for a sufficiently long time, and to have sufficiently pure conditions during growth. Growth mechanism has not been studied but it is clear from the highly perfect silicon found in the tetrahedra enclosed

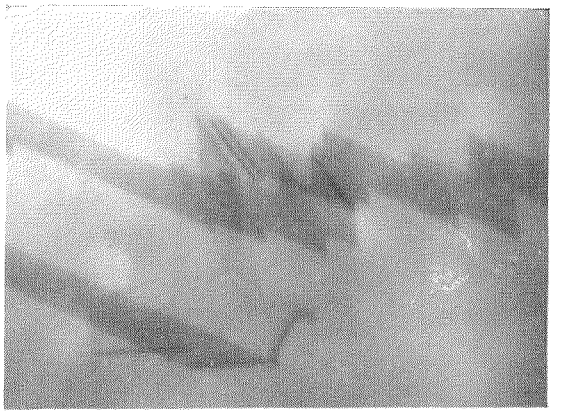


Figure 13—Transmission electron micrograph showing pinning of a dislocation. (× 45 000)

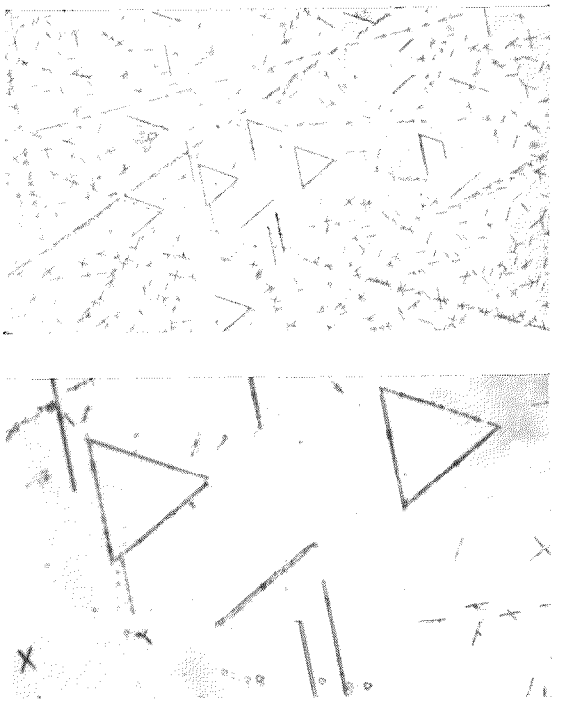


Figure 14—Preferential copper precipitation at stacking faults. Top print is × 300 and bottom is × 900.

by stacking faults that dislocations are not important. It is believed that the stacking faults found in the layers are primarily produced near the start of growth particularly when growing nuclei meet. There appears to be no good reason why layers free from both dislocations and stacking faults cannot be produced. The properties of vapour-grown silicon appear to be substantially the same as melt-grown material; for example, diffusion constants [8] and Hall mobilities [9] are similar.

6. References

1. W. Shockley, "Problems Related to *p-n* Junctions in Silicon," *Solid-State Electronics*, volume 2, pages 35–67; June 1961.

2. W. C. Dash, "Method for Measuring the Thickness of Epitaxial Silicon," *Journal of Applied Physics*, volume 33, pages 2395–2396; July 1962.

3. D. J. D. Thomas and R. C. G. Swann, Technical Memo 366, Standard Telecommunications Laboratories; 1962.

4. H. J. Queisser, "Slip Patterns on Boron-Doped Silicon Surfaces," *Journal of Applied Physics*, volume 32, pages 1776–1780; September 1961.

5. T. H. Schwuttke and H. J. Queisser, "X-Ray Observations of Diffusion-Induced Dislocations in Silicon," *Journal of Applied Physics*, volume 33, pages 1540–1542; April 1962.

6. H. J. Queisser, R. H. Finch, and G. J. Washburn, "Stacking Faults in Epitaxial Silicon," *Journal of Applied Physics*, volume 33, pages 1536–1537; April 1962.

7. W. C. Dash, "Growth and Perfection of Crystals," John Wiley (1958); page 361.

8. C. F. Drake (Private communication.)

9. J. Whittaker (Private communication.)

7. Acknowledgment

It is a pleasure to acknowledge the co-operation of Mr. H. F. Sterling and Mr. R. C. G. Swann, who provided the samples used in this investigation.

K. O. Batsford was born in London on 16 April 1930. He obtained a degree in chemistry at London University in 1951 and in 1954, the degree of Ph.D.

In 1954, he joined the materials division at Standard Telecommunication Laboratories where he was concerned with X-ray and metallographic studies, particularly concerning the crystal perfection of semiconductors. Dr. Batsford is presently engaged in studying material parameters using X-ray, electron-microscope, low-temperature, and infra-red techniques.

D. J. D. Thomas was born in London on 2 June 1927. He obtained a degree in metallurgy at the University College of South Wales and Monmouthshire in 1946.

In 1949, after working in the tin-plate industry, he joined Standard Telecommunication Laboratories, where he is currently engaged with problems connected with semiconductor materials, such as, the purification and growth of germanium crystals, relationship between crystal defects and growth conditions in germanium and silicon, the behaviour of impurities in semiconductors, and defects in vapour-grown materials.

Silicon Epitaxial Planar Transistors Part 1—Appraisal

B. D. MILLS

Standard Telephones and Cables Limited; Footscray, Sidcup, Kent, England

Three major technologies are competing today for producing very-high-frequency transistors, which are described by the following designations:—

(A) Micro alloy diffused transistors (MADT)

(B) Post alloy diffused transistors (PADT)

(C) Diffused transistors.

(A) and (B) are available only in germanium, but the diffused transistor is available also in silicon. In silicon, the marvellous property of silicon dioxide to provide a selective mask against the diffusion of some elements, allows double-diffusion techniques to be used with localized emitter structures, thus giving low input capacitance. A cross-section of a diffused silicon transistor is shown in Figure 1. In germanium, this oxide property is not available, and the transistors have diffused bases but alloyed emitters.

Diffused mesa transistors have been available for many years and have represented a significant advance over preceding technologies. However, they had several shortcomings and various manufacturing difficulties.

1. Diffused Mesa Transistors

Table 1 summarizes the limiting characteristics of representative diffused transistors that were commercially available in January 1961. In considering these figures it must be appreciated that not all these characteristics were available simultaneously in a single transistor. It is unfortunately true that even now it is not possible to make a 150-volt 10-ampere 750-megacycle-per-second transistor with an input capacitance of 2 picofarads. This table is meant more as a guide to the circuit designer to help in deciding whether suitable transistors are available for his particular application. For instance, if the line voltage in an equipment is 50 volts, the designer will be obliged to use a silicon diffused transistor. Alternatively, the circuit designer may not have realized that current ratings up to 10 amperes in silicon were available; and so on. It will be noticed that in voltage, current, power dissipation, and leakage current, silicon offered a greater flexibility than germanium, whereas in frequency there was little to choose between them. Advances have since been made in the frequency capability of both silicon and germanium, and transistors that can operate at 1 gigacycle per second are available in silicon and at 2 to 3 gigacycles per second in germanium.

TABLE 1

LIMITATIONS OF DIFFUSED MESA TRANSISTORS AS OF JANUARY 1961

Characteristics	Measurement Unit	Diffused Silicon	Diffused Germanium
V_{CB0}	Volts	150	40
V_{EB0}	Volts	12	4
I_C, Mean Maximum	Amperes	10	0·5
Power Dissipation	Watts	125	6
f_{hb}	Megacycles per Second	750	750
C_{0b}	Picofarads	2·0 – 3·5	1·2–1·5
I_{CB0}	Microamperes	0·01–0·1	1

It will be noticed that in Table 1 no values are listed for saturation voltage V_{CES}, which depends greatly on the current levels involved. To quote a typical figure, a circuit required to switch 0·5 ampere would involve a bottoming

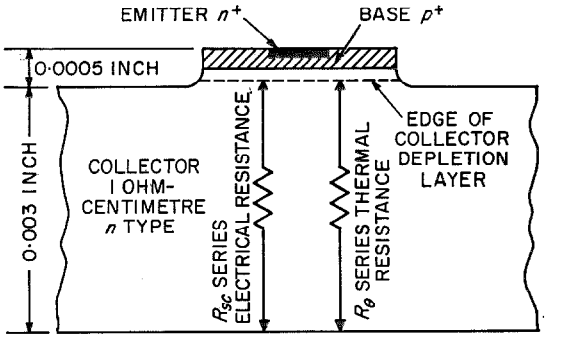
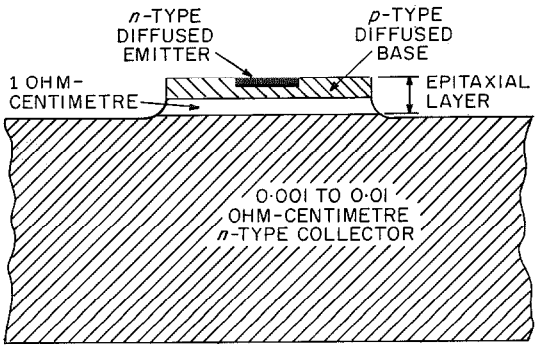


Figure 1—Double-diffused silicon mesa transistor.

Figure 2—Cross-section of an epitaxial double-diffused silicon transistor.



voltage in the range of 1·5 to 3·5 volts. Similarly, switching times are not given as they depend greatly both on the current levels involved and on the actual circuit used. A given transistor may produce quite different storage times in two different circuits. It is unfortunate that it is not possible to present realistic figures for these two characteristics of diffused transistors, because undoubtedly they represent the two major shortcomings of normal diffused mesa transistors. There are three other minor shortcomings of diffused transistors; firstly, they offer disappointing noise figures; secondly, they have low current gain at low currents; and, finally, they often show a spread in characteristics, particularly in output capacitance and in ON voltage (V_{BES}). The two major shortcomings have been overcome by the epitaxial technology and gold diffusion, and the other three by the planar technology which, in so doing, has introduced many other advantages, not least of which is its very high order of reliability.

2. Epitaxial Transistors

The reason for the high saturation voltage can be seen in Figure 1. This shows a cross-section of a double-diffused silicon mesa transistor, but the argument applies equally well to the germanium mesa transistor. The base layer in a diffused transistor may be some 4 or 5 microns below the surface, and at maximum applied collector voltage the depletion layer may extend into the collector region a further 4 or 5 microns. Thus, the operating part of a transistor is in fact only about 9 or 10 microns thick; that is, less than 0·0005 inch. Unfortunately, silicon being very brittle, it is not possible to handle large pieces of silicon of this thickness in mass production. Indeed, for mass production of diffused transistors, the minimum practical thickness of semiconductor is 0·003 inch (76 microns). To obtain the desirable characteristics of high collector breakdown voltage and low output capacitance, a rather high resistivity (1 to 6 ohm-centimetres) is required at the collector-to-base junction. The unwanted thickness of silicon introduces a series electric resistance and a series thermal resistance, both of which degrade the inherent properties of the mesa transistor. This series electric resistance is the source of the high saturation voltage of normal diffused mesa transistors.

The epitaxial transistor has provided a brilliant solution to this problem. An epitaxial layer is one that has the same crystalline orientation as the substrate on which it is grown. If silicon is exposed at 1200 degrees centigrade to hydrogen containing 1 percent of silicon tetrachloride, the tetrachloride will decompose and silicon will be deposited on the silicon substrate. At these temperatures the silicon atoms condensing on the surface of the substrate are quite mobile and are able to take up the orientation of the substrate lattice. By adding minute amounts of phosphorus trichloride or boron tribromide to the silicon tetrachloride it is possible to control the resistivity of the deposited layer.

For the epitaxial transistor a high-resistivity (1-to-6-ohm-centimetre) layer of silicon about 12 microns thick is deposited epitaxially on a low-resistivity (0·001-to-0·01-ohm-centimetre) piece of silicon some 0·005 inch (127 microns) thick. A transistor is then fabricated in the normal way by oxide masking and double-diffusion techniques, as shown in Figure 2.

The advantages of the epitaxial transistor are

(A) Low saturation resistance

(B) Saturation resistance independent of temperature

(C) Increased linearity of characteristics

(D) Reduced minority-carrier storage time.

The improved saturation characteristics are apparent from Figure 3, which shows the $I_C - V_{CE}$ characteristics of two 100-megacycle-per-second silicon power transistors, identical except that one used epitaxial material.

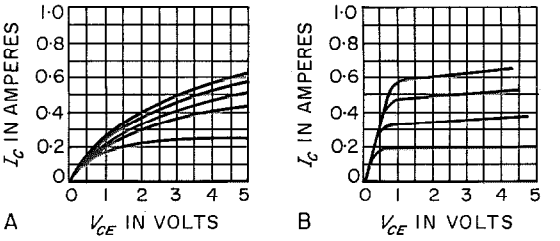
The implications of the lower saturation voltages will be obvious to circuit engineers. In switching circuits greater efficiencies are achievable, leading to lower dissipations; or for a given transistor design, greater power-handling capacity. In amplifier and oscillator applications, greater voltage swings are possible, giving increased output powers to the load. An amplifier using two non-epitaxial *BLY10* transistors in push-pull from an 18-volt supply gave 4 watts output power at 5 megacycles per second. By merely substituting epitaxial transistors, the output power was increased to 9 watts at 55-per-cent efficiency. By redesigning the circuit slightly to take advantage of the lower saturation resistance, an output power of 10 watts was obtained at 5 megacycles per second.

The second advantage of epitaxial transistors shown above derives from the fact that the collector series component of total saturation resistance has been virtually eliminated by the epitaxial construction. The actual degree of improvement depends very much on the resistivity of the collector material used in the non-epitaxial transistor; but one typical comparison shows an increase in saturation resistance between 25 and 150 degrees centigrade of 2·15 times for a *BLY10* non-epitaxial unit and only 1·15 times for a *BUY10* epitaxial transistor.

Thus, on the epitaxial transistor, the saturation resistance increased only 15 percent from room temperature to 150 degrees centigrade.

The improved linearity (and higher gain) is also obvious from Figure 3. Thus, with epitaxial transistors lower distortion levels can be obtained, which make them attractive for amplitude-modulated transmitter applications.

Figure 3—Typical characteristic curves for two silicon power transistors, that at the right being for a *BUY11* epitaxial design and that at the left is for a *BLY11* non-epitaxial unit.



Finally, the improvement in switching performance offered by epitaxial transistors is due to the location of a low-resistivity layer close to the collector-to-base junction. In saturated switching circuits the collector is forward-biased in the ON condition, and at turn-off the minority carriers that have been injected from the base into the collector have to be removed before the transistor will switch off. Circuit conditions limit the rate at which these carriers can be removed, and the turn-off time is dependent to a large extent on the lifetime of the minority carriers in this region. The presence of a very-low-resistivity region near the collector junction limits the number and lifetime of the carriers that are injected, thus reducing the storage time in switching circuits.

In a specific circuit, a low-level switching transistor gave an average storage time t_s of 550 nanoseconds. The epitaxial version of this transistor gave a t_s value of 100 nanoseconds.

3. Gold-Doped Transistors

To switching-circuit engineers, the storage times quoted above will not seem very attractive.

There is, fortunately, another technology available with the diffused transistor. This is gold doping, by which it is possible to control

the lifetime of minority carriers. Gold is diffused into the base and collector regions of the transistor; this reduces the storage time in switching circuits. For example, in the low-level epitaxial transistor quoted above, the storage time of 100 nanoseconds was reduced to 10 nanoseconds and less. This speed should be acceptable for most switching circuits.

4. Silicon Planar Transistors

In the mesa transistor, the area of the collector-to-base junction is defined by etching a mesa. The planar transistor is made by countersinking the base into the silicon, using photolithographic techniques to establish a suitable oxide mask. The reason for the high reliability and the other attractive properties of the silicon planar transistor will be better understood by considering Figure 4, which shows the most significant steps in the fabrication of the planar transistor. It should be appreciated that the transistor shown in this figure is just one of up to 1200 similar transistors formed on a single slice of silicon about 1 inch (25 millimetres) in diameter.

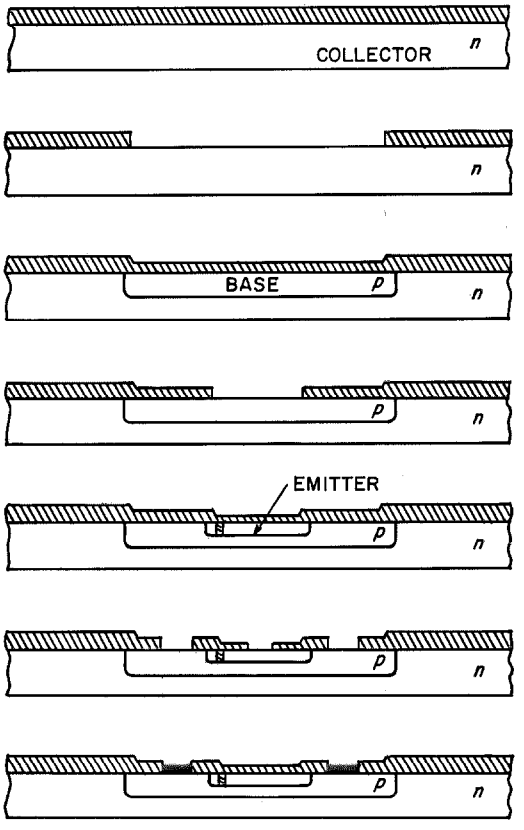


Figure 4—Steps in fabricating planar transistors.

The first step is to grow an oxide layer of 0·7-micron thickness on the silicon surface under scrupulously clean conditions. The area of the base is then defined by a photolithographic process, and a window etched in the silicon dioxide. This process is very precise and repeatable. The slice is now exposed at high temperatures to a boron-rich atmosphere. Silicon dioxide masks the diffusion of boron and, hence, a *p*-type layer is formed only where the window was etched in the silicon dioxide. Fortunately, boron diffuses isotropically in silicon, so that lateral diffusion of boron takes place as well as diffusion normal to the surface. This might not appear so from Figure 4, which is not drawn to scale in order to show all the information in one figure. Thus, the junction is not formed opposite the edge of the layers, but 4 or 5 microns *underneath* the oxide layer. This process gives control of the surface states at the *p-n* junction, and it is from this control that all the advantages of the planar transistor derive. Once a *p-n* junction is formed, an electric field exists across it, and if this junction is exposed it will attract ionic contaminants to it, thus degrading its properties. In the planar transistor the sealed junction prevents this from happening, and the properties of the junction, once made, remain unchanged throughout the fabrication and life of the transistor. During boron diffusion more oxide is grown on the surface, so that we are able to repeat the photolithography and diffusion processes with phosphorus to produce an *n*-type emitter. Note that the emitter junctions are also underneath the oxide layer. Finally, during phosphorus diffusion more oxide is grown to give a completely sealed transistor. Windows are etched in appropriate places in this oxide to allow contacts to be made to the operative areas of the transistor.

This transistor is now complete; it has its final characteristics and requires no etching or other treatment to improve its characteristics.

A 400-megacycle-per-second silicon planar transistor is shown in Figure 5. Its diameter is 0·007 inch (175 microns).

The advantages of planar transistors are summarized below:—

(**A**) Very-low leakage currents

(**B**) High gain at low currents

(**C**) Low noise figure

(**D**) Very-consistent characteristics

(**E**) Very-stable characteristics

(**F**) Very-high order of reliability.

As a result of the sealed junctions, the leakage currents in silicon planar transistors are very low indeed. On a 10-watt power transistor, which is rated at 1·5 amperes peak current, the typical leakage currents are in the range 1 to 3 nanoamperes. Low-level devices give typical leakage currents in the range 0·1 to 1 nanoampere. This makes the planar transistor unquestionably the most suitable for direct-current amplifiers.

Due to surface effects, mesa transistors have very-low gains at currents in the microampere range. At 1 microampere, the small-signal gain might be 0·3 to 3. With planar transistors, gains as high as 100 at 1 microampere have been achieved, and transistors are sold with minimum gain of 60 at 100 microamperes. This feature of planar transistors could be important in compact designs where power densities must be minimized.

Planar transistors offer considerably lower noise figures than equivalent mesa transistors; this results from the control of the surface states of the *p-n* junctions. Low-frequency noise figures in the range 3 to 10 decibels at a few hundred microamperes are typical. The low-frequency noise figure is current-dependent and increases with increasing collector current beyond 1 milliampere. An interesting experiment is to measure the noise on a planar transistor, and then remove the oxide with a weak acid. The noise figure increases by 8 to 12 decibels. The low noise figure is an advantage in input stages of direct-current amplifiers, as well as in many other circuits. At 1 megacycle per second, noise figures of 1 to 2 decibels are available with planar transistors.

The greater consistency of characteristics offered by planar transistors comes from two sources. Firstly, the base area, which defines the output capacitance, is determined with almost mathematical precision by a photolithographic process. Secondly, and by far the more important, is the fact that there is no final etching required with this transistor. Before encapsulation of any exposed-junction transistor, final etching is necessary to remove the ionic contaminants that have been attracted by the electric field during the

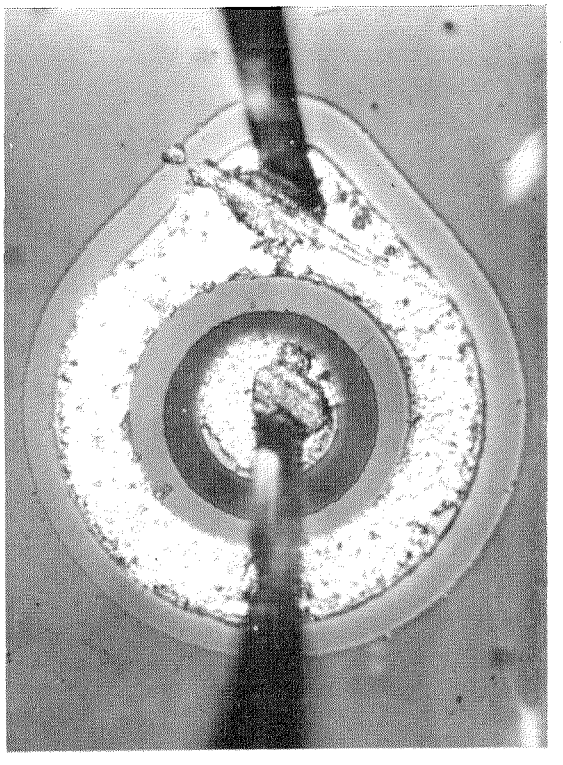


Figure 5—A 400-megacycle-per-second silicon planar transistor.

fabrication process. This restores the original characteristics. Variable amounts of such etching can affect the area of the mesa and, hence, the output capacitance. An even-more-serious effect of this etching is on the ON voltage or V_{BES}. The V_{BES} is determined mainly by the extrinsic base resistance which, in a diffused transistor, depends on the resistance between the aluminium base contacts and the operative area of the base. Base current flows through the path of least resistance along the surface to the edge of the emitter and thence to the operative part of the base. The surface areas are of very-low resistivity due to the diffusion profile, and when the collector-base junction is etched, the surface layers are attacked very rapidly; this can result in a considerable spread in extrinsic base resistance and, hence, in the ON voltage.

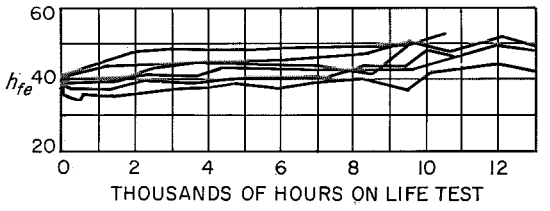


Figure 6—Values of h_{fe} for 6 *BLY11* silicon planar transistors measured at 1 kilocycle per second with $V_{CE} = +9$ volts, $I_C = +20$ milliamperes, and 3 watts dissipation.

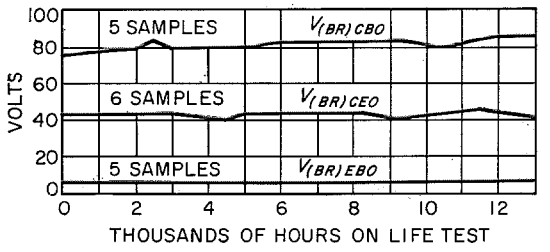


Figure 7—Type *BLY11* silicon planar transistors life tested at 3 watts dissipation. $V_{(BR)CB0}$ was measured with $I_C = +1$ milliampere, $V_{(BR)CE0}$ with $I_C = +2$ milliamperes and $I_B = 0$, and $V_{(BR)EB0}$ with $I_C = 0$ and $I_E = +1$ milliampere.

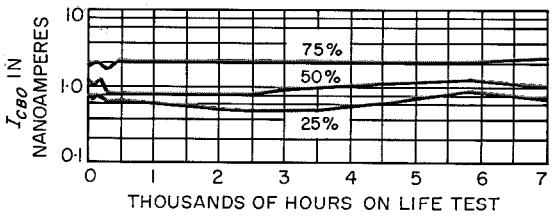


Figure 8—Unencapsulated silicon planar transistors, type *BFY15*, stored at 175 degrees centigrade and measured at $V_{CB} = +9$ volts and $I_E = 0$. The percentage of the 24 samples on test having measured values not exceeding those of the curve is indicated on each curve.

The planar transistor offers a high order of characteristic stability due to the control of the surface states at the emitter-base and collector-base junctions. Figures 6 and 7 show stability results on some very early experimental *BLY11* silicon planar transistors; h_{fe} characteristics may leave something to be desired, but it will be seen from later results that this performance has been considerably improved. These transistors have been dissipating 3 watts in free air throughout their operating life. A sample of 6 units showed an unvarying maximum value of I_{CB0} of 0·003 microampere over a life test of 13 000 hours with V_{CB} of 9 volts, I_E of zero, and dissipation of 3 watts.

The very-high order of reliability offered by planar transistors is not dependent on the quality of their encapsulation or the control of the ambient during life, but is inherent in the design and fabrication of the transistor. Firstly, the key areas of the transistor, the *p-n* junctions, are sealed beneath an impervious layer of silicon dioxide. That this layer is impervious can be illustrated by the fact that unencapsulated planar transistors can be operated under water without any change in their gain or breakdown voltage characteristics! Secondly, a high order of mechanical reliability is introduced by the absence of etching during the fabrication process. Very strong acids are necessary to etch silicon junctions so that while clean-up etching is in progress in mesa transistors, chemical attack also proceeds on the metal parts. In particular, the gold-alumin-

ium-silicon bonds of the transistor can be severely attacked and weakened. Planar transistors are sufficiently new so that reliability results are of limited duration in terms of total operational hours. As further evidence of this reliability, however, storage on these transistors for 7000 hours at room temperature and at 150 degrees centigrade have produced no changes in their characteristics. Storage life tests at 175 degrees centigrade have passed 8000 hours to date without any appreciable change in gain, breakdown, or leakage characteristics.

A series of experiments has established that the reliability of the transistor is inherent in the device and not dependent on its encapsulation. Unencapsulated transistors have been stored at room temperature and at 115, 135, and 175 degrees centigrade. Storage at room temperature has reached some 10 000 hours without any apparent deterioration in gain, breakdown voltage, or leakage current. Storage tests under the other three temperatures have now reached 7000 hours and similarly show no apparent changes in characteristics. Figures 8, 9, and 10 show the results at 175 degrees centigrade, which are typical of all temperatures. We have also subjected unencapsulated transistors to a range of severe environmental tests that include humidity cycling, salt-fog tests, and boiling in deionized water for prolonged periods. Unencapsulated transistors have successfully withstood 100 cycles of the standard United Kingdom Government humidity cycling test. Rust from the metal parts of the mounting completely covers the transistor and causes an increase in the total leakage current; but the breakdown voltage and gain remain unaffected. Similar results are achieved with the salt-fog test and the boiling-water test. The last two tests have been applied for over 500 hours without producing any change in the gain or breakdown voltage. Figures 11, 12, and 13 show the results on the salt-fog test; these results are typical of the others. After about 550 hours the metal parts of the transistor including the leads have also rusted away, and it is this fact that brings the test to an end, *rather than any deterioration in the transistor*. To prove that the increased leakage currents were due only

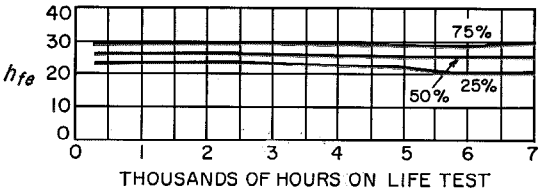


Figure 9—Unencapsulated *BFY15* transistors stored at 175 degrees centigrade and measured at $V_{CE} = +9$ volts and $I_C = +20$ milliamperes. Percentage of the 24 samples not exceeding the values of each curve are indicated.

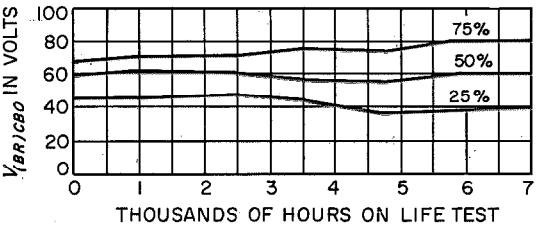


Figure 10—Unencapsulated *BFY15* transistors stored at 175 degrees centigrade and measured at $I_C = +10$ microamperes and $I_E = 0$. Percentage of the 25 samples not exceeding the values of each curve are indicated.

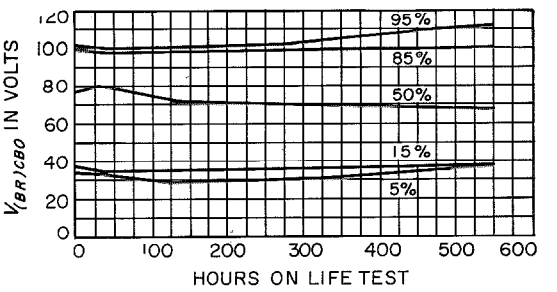


Figure 11—Unencapsulated silicon planar transistors, type *BFY16*, subjected to salt spray at room temperature. $V_{(BR)BC0}$ was measured at $I_C = +10$ microamperes and $I_E = 0$. Percentage of 24 samples having measured values not exceeding those of the curves is indicated on each curve.

to the parallel path of the rust, the transistors were dipped in a 50-percent hydrochloric-acid solution that removed the rust and returned the leakage currents to their initial values. This can be seen clearly in Figure 13.

Similarly, unencapsulated transistors have been boiled in saturated brine solutions for several hours without any deleterious effects.

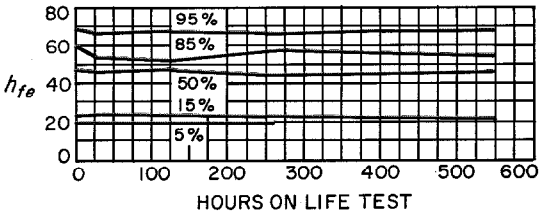


Figure 12—Unencapsulated *BFY16* transistors subjected to salt spray at room temperature. Measurements of h_{fe} were made at $V_{CE} = +9$ volts and $I_C = +20$ milliamperes.

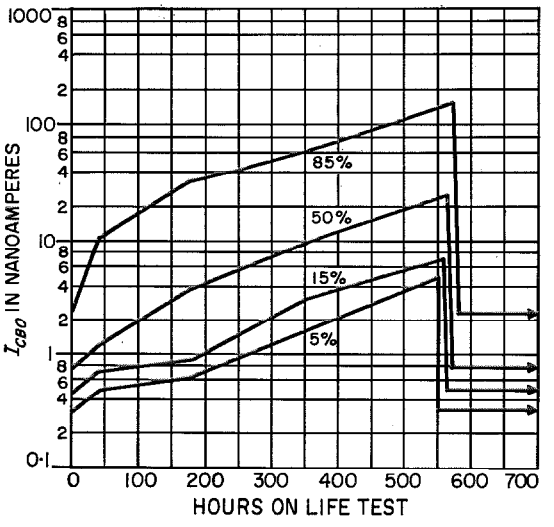


Figure 13—Unencapsulated *BFY16* transistors subjected to salt spray at room temperature. Measurements of I_{CB0} were made at $V_{CB} = +9$ volts and $I_E = +0$. At 550 hours, the corrosion products, which included the leads, were removed in hydrochloric acid and the initial values reappeared indicating that the deterioration was a superficial effect of corrosion.

Thermal-shock tests have also been carried out on these transistors to ensure that no troubles could arise from possible differences in expansion coefficients between the silicon and the silicon dioxide layer. Transistors have been stabilized at 150 degrees centigrade and then plunged into liquids at −70 degrees centigrade. This process has been repeated many times without any apparent effect on the transistor characteristics.

Finally, a whole series of sensitive ambient tests has been conducted on these transistors to determine any trace of surface instability. No such instability was found.

5. Silicon Epitaxial Planar Transistors

The planar design can, of course, be combined with the epitaxial processing and the gold-doping technology to produce the silicon epitaxial planar transistor. This is probably the most versatile and reliable transistor presently available. It may be thought that such reliable and versatile transistors would of necessity be expensive. This is completely invalid, as diffused transistors are made by production techniques that render them inherently cheap. Firstly, all the diffusion and photolithographic processes are performed simultaneously on up to 1200 transistors. Secondly, with the planar process it is possible to test the transistors at every stage in manufacture so that a poor slice can be thrown away before large costs have been involved. Finally, each individual transistor die can be fully tested before assembly, thus ensuring high assembly yields. Thus, silicon epitaxial planar transistors have many advantages from a manufacturing point of view, as well as from the technical aspect. It is confidently predicted that their prices will reach the level of germanium alloy transistors in due course.

6. Conclusions

The epitaxial planar technology is a very versatile one, which can be used over a wide range

of frequency, current, and power levels. Gigacycle-per-second amplifiers and 20-watt 10-ampere devices can both be made by planar techniques. It is felt that silicon planar epitaxial transistors can fulfil a large percentage of all transistor applications, with an order of reliability that has not previously been achieved.

This technology is now also well established in solid-state circuits, where its flexibility and reliability have led to significant advances.

Bernard Douglas Mills was born in Charters Towers, Queensland, Australia, in 1928. He received the degree of Bachelor of Electrical Engineering in 1950 and Bachelor of Science in 1951 from Queensland University in Brisbane.

After 18 months in the power industry, he served as a development engineer on radio receiving valves.

He joined Standard Telephones and Cables in 1957, and was engaged in work on semiconductor devices. He is now chief engineer of the Transistor Division.

Silicon Epitaxial Planar Transistors
Part 2—Characteristics

J. BICKLEY
Standard Telephones and Cables Limited; Footscray, Sidcup, Kent, England

The range of silicon planar transistors manufactured by Standard Telephones and Cables contains more than 15 codes, and it must first be made clear why this number has been brought into existence, and why there is not a universal transistor.

Any such device represents a compromise between cost and performance over the wide range of applications for which transistors are used. Although a low-cost transistor can be made to give good performance over a wide range of applications, for example, the *BSY27*, better performance over a narrower range of applications, probably at a higher cost, can always be obtained. Consequently, although there are general-purpose transistors, there is never likely to be a universal transistor.

The main characteristics of the range of silicon planar transistors are given in Table 1.

1. Direct-Current

Figure 1 shows the variation of large-signal current gain h_{FE} with collector current I_C for representative specimens from each family.

The highest gain device is the *BFY19*, the 400-megacycle-per-second non-epitaxial amplifier. However, the gain falls sharply at collector currents beyond 30 milliamperes. The epitaxial version *BSY29*, on the other hand, has a lower maximum gain, but maintains its gain at a useful level to 200 milliamperes. It has the widest range of useful gain of all these devices.

The *BSY27* has a peak gain similar to the *BSY29* but, being of larger area, the peak occurs at a higher value of collector current. The non-epitaxial version of the *BSY27*, the *BFY18*, peaks at the same collector current, but as in the case of the *BFY19*, the gain falls sharply at higher currents due to the collector series resistance.

Figure 1 also shows the excellent gains of the large-area *BUY11* and *BSY25* at collector currents up to the peak rating of 1·50 amperes.

TABLE 1
CHARACTERISTICS OF CODED SILICON TRANSISTORS

Family	Type	Encapsulation	Dissipation in Watts	Typical f_T in Mega-cycles per Second	Typical h_{FE} or h_{fe}	Structure	Typical Application
A	*BLY10* *BLY11*	TO-3 TO-3	10 With Heat Sink	100 150	30 60	Planar	High-Power High-Frequency Oscillators and Amplifiers
	BFY15 *BFY16*	TO-5 TO-5	0·6 in Free Air	100 150	30 60	Planar	Medium-High-Frequency Power Oscillators and Amplifiers
	BUY10 *BUY11*	TO-3 TO-3	10 With Heat Sink	100 150	30 60	Planar Epitaxial	High-Level Switches, High-Efficiency Power Oscillators and Amplifiers
	BSY24 *BSY25*	TO-5 TO-5	0·6 in Free Air	100 150	30 60	Planar Epitaxial	High-Current Switches, Medium-Power Oscillators and Amplifiers
B	*BFY17* *BFY18*	TO-5 TO-18	0·6 in Free Air 0·3 in Free Air	300 300	60 60	Planar	High-Frequency Amplifiers
	BSY26 *BSY27*	TO-18 TO-18	0·3 in Free Air	300 300	30 60	Planar Epitaxial	General-Purpose Fast Switches and Amplifiers
C	*BFY19*	TO-18	0·3 in Free Air	400	100	Planar	Very-High-Frequency Amplifiers
	BSY28 *BSY29*	TO-18 TO-18	0·3 in Free Air	400 400	30 60	Planar Epitaxial	Very-Fast Switches

The curves for the non-epitaxial versions of these devices have not been shown on the diagram to avoid confusion, but their drop in gain occurs at currents beyond 0·50 ampere.

Both large- and small-signal current gain rise with temperature throughout the range from −55 to +175 degrees centigrade. The rate of rise varies somewhat for different types of device, a temperature rise between 100 and 150 degrees centigrade being needed to double the gain.

The advantages of epitaxy are shown most markedly in the variation of V_{CES} with collector current. This is shown in Figure 2 for a typical sample of each family. These curves are for a ratio of I_c/I_B of 10. At low currents V_{CES} varies only slightly, but beyond a certain collector current, which is dependent on the area of the active element and its collector structure and resistivity, the V_{CES} rises as a function of current. It is interesting to note that for the three non-epitaxial types, V_{CES} rises linearly with collector current; whereas for the epitaxial devices V_{CES} rises approximately with $I_C^{\frac{1}{2}}$. This difference is probably due to conductivity modulation of the epitaxial layer in the collector region, stemming from the much-higher current density for the collector currents at which V_{CES} begins to rise with current.

Figure 3 shows the variation of V_{BE} and V_{CES} with temperature. V_{BE} falls at 1·85 millivolts per degree centigrade for both epitaxial and non-epitaxial devices, both absolutely and as a percentage of the room-temperature value.

2. Alternating-Current

The low-frequency gain h_{fe} reaches a maximum at a value of I_C determined roughly by the area of the device in question. In the case of the non-epitaxial devices, for instance, the maximum occurs at 10 milliamperes in the case of the *BFY17* and *18*, and 5 milliamperes for the *BFY19*, compared with 80 milliamperes for the power transistors *BLY10* and *11*. For the smaller-area devices the gain remains at a

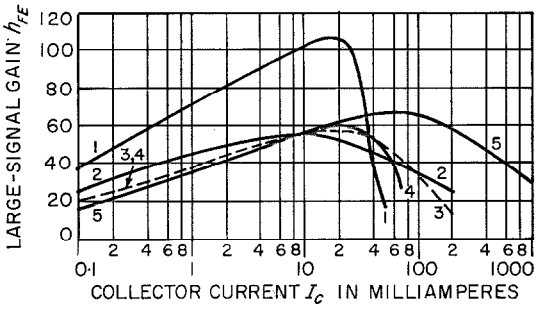


Figure 1—Typical variations of large-signal current gain h_{FE} as a function of collector current I_C. The type of transistor and the measurement V_{CE} (in volts) are as follows: curve 1, *BFY19* (6); curve 2, *BSY29* (2); curve 3, *BSY27* (2); curve 4, *BFY17/BFY18* (6); and curve 5, *BUY11/BSY25* (2).

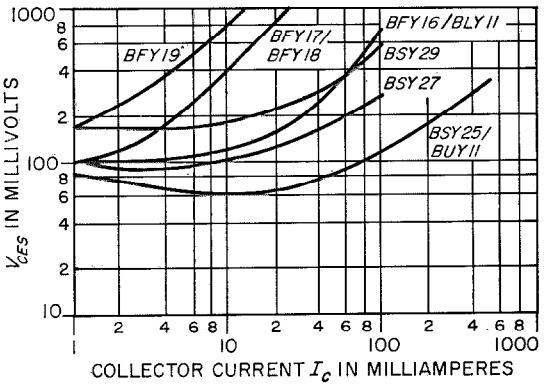


Figure 2—Typical variations of the collector-emitter saturation voltage V_{CES} with collector current I_C for $I_C/I_B = 10$.

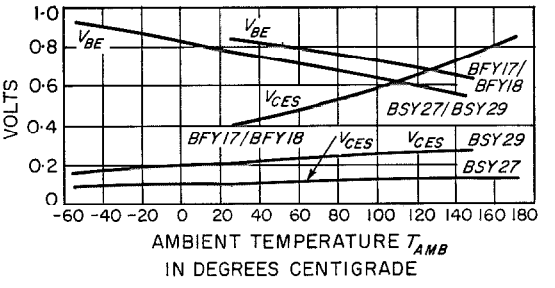


Figure 3—Typical variations of input voltage V_{BE} and collector-emitter saturation voltage V_{CES} with ambient temperature for $I_C = 10$ milliamperes and $I_C/I_B = 10$.

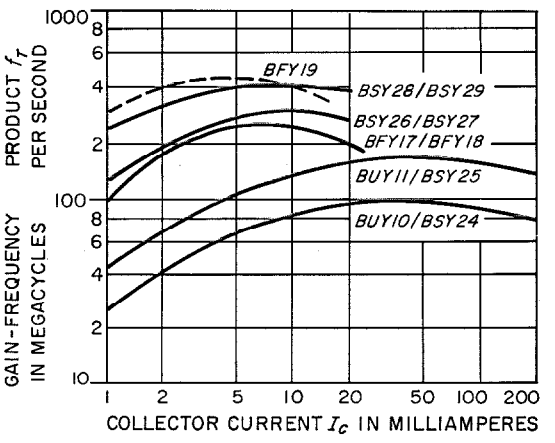


Figure 4—Typical variations of the common-emitter gain-frequency product f_T with collector current I_C.

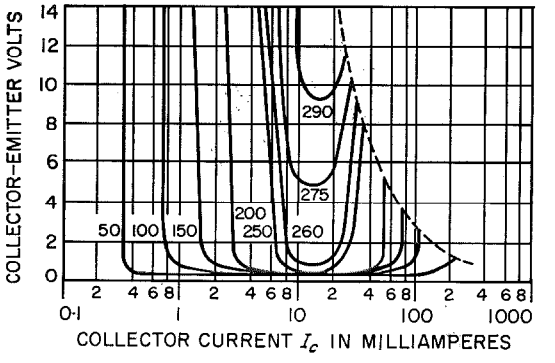


Figure 5—Typical variation for epitaxial *BSY27* of gain-frequency product f_T witsh collector current I_C and collector-emitter voltage V_{CE}. The values of f_T in megacycles per second are ind[1]cated for each curve.

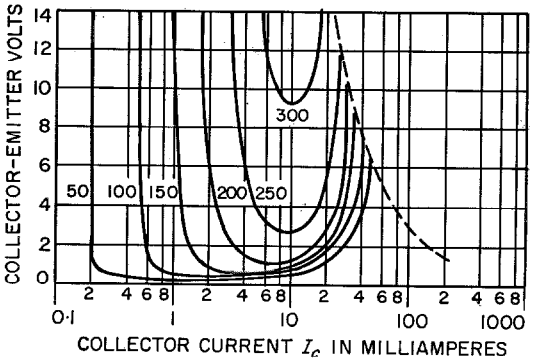


Figure 6—Typical variations for non-epitaxial *BFY17* of gain-frequency product f_T with collector current I_C, and collector-emitter voltage V_{CE}. The values of f_T in megacycles per second are indicated for each curve.

useful value for collector currents down to the microampere region.

The variation of the gain-frequency product f_T as a function of collector current is shown in Figure 4 for representative samples of the three main groups of devices. As in the case of low-frequency gain, f_T reaches a maximum value at a current that becomes smaller for smaller-area devices. The effect of epitaxy is to improve the value of f_T at high currents by preventing the collector-base junction voltage from falling at high currents. However, it can be seen from the curves that the improvement is slight at currents around 5 milliamperes, which are commonly used for small-signal amplifiers.

The effect on the f_T of a particular device of varying the bias conditions can most conveniently be illustrated by plotting contours of constant f_T in the I_C, V_{CE} plane. This has been done in Figures 5 and 6 for the *BSY27* and *BFY17*. These two devices have basically the same structure, but the *BSY27* is epitaxial, unlike the *BFY17*. It will be noticed that the contours of the epitaxial device are more "flat bottomed" than those of the non-epitaxial, that is, the f_T of the epitaxial device varies less with current near its peak value for a given collector voltage and maintains its f_T more nearly constant to lower voltage levels than the non-epitaxial device for currents above 1 milliampere.

The effect of epitaxy is much more marked at high collector currents. For the epitaxial device the contours are nearly symmetrical on either side of the peak f_T axis, but for the non-epitaxial device the contours are very much compressed and distorted towards the higher-voltage and lower-current region. This implies an inferior switching performance at high currents for the non-epitaxial device, even in a non-saturating circuit.

The variations of the common-base output capacitance C_{OB} at zero emitter current with collector voltage for the three basic families of devices are shown in Figure 7. At voltages high enough for the built-in voltage to be

neglected, the capacitance is approximately inversely proportional to the cube root of the applied voltage, as would be expected for a graded junction. The measured capacitance includes about 0·5 picofarad due to the stray capacitances inside the case.

The emitter depletion capacitance C_{Te} is important in switching applications because it influences the delay time, that is, the time between the application of a drive voltage to the base of the transistor and the beginning of the rise in collector current.

In amplifier applications the emitter depletion capacitance causes the fall in f_T at low emitter currents. However, in this case it is the value of C_{Te} when the emitter junction is forward biased that is required, and this cannot be measured directly.

For individual devices an effective value of C_{Te} appropriate to this application can be deduced from measurements of the variation of f_T with emitter current; but for both applications the value of C_{Te} at low reverse bias is useful for comparison purposes.

Typical values of C_{Te} at an emitter-base reverse bias of 0·5 volt for the three main families are: *BSY25* family, 45 picofarads; *BSY27* family, 6·5 picofarads; and *BSY29* family, 3·2 picofarads.

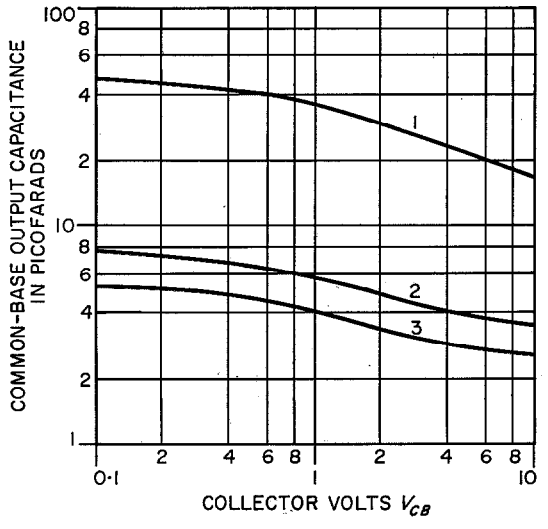


Figure 7—Typical variations of common-base output capacitance C_{OB} with collector-base voltage V_{CB} for $I_E = 0$. Curve 1 is *BSY24/25*, *BFY15/16*, *BUY10/11*, and *BLY10/11*. Curve 2 is for *BSY26/27* and *BFY17/18*. Curve 3 is for *BSY28/29* and *BFY19*.

3. Switching Performance

The differences between devices can be seen by comparing measurements of the times for delay, rise, storage delay, and fall in a particular circuit. Typical results are shown in Table 2.

TABLE 2
TYPICAL SWITCHING TIMES OF EPITAXIAL DEVICES

Times in Nanoseconds	*BSY24*	*BSY25*	*BSY26*	*BSY27*	*BSY28*	*BSY29*
t_d	9	7	8	9	7	7
t_r	10	9	12	10	7	5
t_{sd}	350	310	15	19	8	7
t_f	25	15	16	16	14	13
I_C in Milliamperes	100	100	10	10	10	10

4. Conclusions

This survey has been too brief to describe more than a few of the important parameters of silicon transistors. It is hoped, however, that it will help circuit designers to find the type most suited to their requirements.

J. Bickley received an honours degree in electrical engineering from Queen Mary College of the University of London in 1954. After three years of research in liquid-dielectric breakdown, he joined the Royal Naval Scientific Service and worked on measuring techniques for high-frequency silicon mesa transistors at the Services Electronic Research Laboratories at Baldock.

He joined the Transistor Division of Standard Telephones and Cables in May 1960 to develop measuring techniques for the diffused silicon high-frequency transistors then being developed. At present he is in charge of the device-characterization laboratory.

Mixer Diodes for 6 Gigacycles Per Second

J. B. SETCHFIELD

Standard Telecommunication Laboratories Limited; Harlow, Essex, England

1. Introduction

It is almost universal practice in microwave communication systems to amplify the signal to be passed on from one repeater to the next at a lower more-convenient frequency. This involves superheterodyne frequency conversion, first down to the chosen intermediate frequency and then, after amplification, up to the original or to a new microwave frequency. The accepted non-linear elements used for conversion in this frequency range are semiconductor diodes. For down-conversion the non-linear-resistance property of these diodes is invariably used while for up-conversion there is a choice between non-linear-resistance and non-linear-capacitance action. Each has its own merits, and while the latter might be the popular choice today on the grounds of greater conversion efficiency, the former can be designed to exhibit certain worthwhile advantages that can outweigh the disadvantage of a slightly higher conversion loss. In the present case non-linear resistance action for the up-converter was chosen to achieve a controlled intermediate-frequency limiting level.

Below are detailed the target specifications for the two diodes, a low-level down-converter and a medium-level up-converter. It will be seen that in addition to qualities that compare with the best mixer diodes available, certain unusual features were demanded. For the down-converter a very-low intermediate-frequency impedance was required and for the up-converter an envelope- or amplitude-limiting characteristic had to occur at a prescribed intermediate-frequency drive level.

Further justification for the choice of non-linear-resistance action for the up-converter lay in the request for an exceptionally low level of harmonics of the intermediate frequency in the super-high-frequency band. It is well known that variable-capacitance action results in the efficient production of harmonics whereas variable-resistance diodes, besides being inherently less efficient in this respect, can be designed to suppress this tendency still further. There may be ways of obtaining similar suppressions in reactance modulators while emphasizing desired modulation products, but so far no exhaustive investigations are known that have explored this in the depth done with resistance modulators.

1.1 Target Specifications

Low-Level Down-Converter

Super-High-Frequency Range = 5·8 to 6·5 Gigacycles Per Second

Local-Oscillator Level for Optimum Noise Figure = 1 Milliwatt

Noise-Temperature Ratio = 1·5, Maximum

Conversion Loss = 4·5 Decibels, Maximum

Intermediate-Frequency Source Impedance at the Above Local-Oscillator Level = 100 to 150 Ohms

Intermediate Frequency = 70 Megacycles Per Second

Burnout Level (Continuous) = + 8 Decibels Referred to 1 Milliwatt, Minimum

Ambient Temperature Range = 10 to 60 Degrees Centigrade

Spread in Radio-Frequency Impedance Match = 20 Per Cent, Maximum

Encapsulation = Hermetic

Medium-Level Up-Converter

Super-High-Frequency Range = 5·8 to 6·5 Gigacycles Per Second

Intermediate-Frequency Impedance of Two Diodes in Parallel and in Reverse Polarity Conducting 8 Milliamperes Each = 100 ± 15 Ohms

Conversion Loss (Super-High to Super-High Frequency) at Super-High-Frequency Input of ±15 Decibels Referred to 1 Milliwatt and Rectified Intermediate-Frequency Current of 8 Milliamperes = 11 Decibels, Maximum

Direct Harmonics of 70-Megacycle-Per-Second Signal at 10 Milliamperes of Rectified Intermediate-Frequency Current = − 120 Decibels Referred to 1 Milliwatt, Maximum

Flatness of Rectification Response from 40 to 100 Megacycles Per Second = 0·5 Decibel with 5-Per-Cent Maximum Spread

Super-High-Frequency Burnout Level with 8 Milliamperes of Rectified Current (Per Diode) = 100 Milliwatts, Minimum

Intermediate-Frequency Burnout Level with ±15 Decibels Referred to 1 Milliwatt Super-High-Frequency Input = 15 Milliamperes of Rectified Current, Minimum

Ambient Temperature Range = 10 to 60 Degrees Centigrade

Envelope Limiting (Intermediate- to Super-High-Frequency Compression of Amplitude Modulation) at 8 Milliamperes Rectified Current = 5 to 1, Minimum

Spread in Radio-Frequency Impedance = 20 Per Cent, Maximum

Encapsulation = Hermetic

2. Low-Level Down-Converter

It is assumed that the reader is familiar with the principles of frequency mixing and conversion and the peculiar requirements at microwave frequencies that dictate the use of the point-contact semiconductor diode.

2.1 Design Parameters

The most-revealing property of any radio receiver is its over-all noise figure, which defines the sensitivity limit. In a superheterodyne down-converter employing a crystal diode the parameters of the diode are related to the receiver over-all noise figure thus:—

$$F_R = L_X(t_x + F_{IF} - 1)$$

where F_R = receiver over-all noise figure, expressed as a ratio.

L_X = conversion loss of the crystal mixer, which is defined as the ratio of the available power at the signal input to the available power at the intermediate-frequency terminals.

t_x = noise-temperature ratio of the crystal diode, which is defined as the ratio of the available noise power output of the diode to that of an equivalent resistor at room temperature.

F_{IF} = noise figure of the intermediate-frequency amplifier, expressed as a ratio.

Conversion loss and noise-temperature ratio are seen to be the significant diode parameters with the former playing the dominant role since F_{IF} is always greater than unity.

2.1.1 *Conversion Loss*

The microwave mixer is usually represented as a time-varying conductance between the radio-frequency input terminals and the intermediate-frequency output terminals. As can be seen from Figure 1, a diagrammatic representation of a crystal diode, this time-varying conductance corresponding to R_B, the barrier resistance, cannot be realized in semiconductor diodes without the accompanying degrading elements C_B, the barrier capacitance and R_S, the bulk series resistance of the inactive part of the semiconductor. It can be said that the conversion loss is determined by the non-linear barrier resistance alone and that the degradation above the theoretical minimum is due to the parasitic elements C_B and R_S.

Considering the effect of R_B alone, the current-voltage relationship for the junction is

$$I = I_o \exp[(qV/kT) - 1].$$

The conversion loss depends on the degree of non-linearity which in turn depends on the voltage swing V and thus is a function of the local-oscillator drive. As the local-oscillator

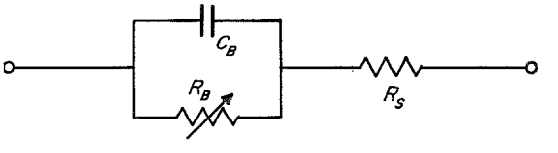


Figure 1—Network of principal elements of a crystal diode.

drive is increased from zero, conversion loss decreases from infinity to a theoretical minimum of 2. When the effects of the parasitics C_B and R_S are added it is seen that for a weak local-oscillator drive the value of R_B is high and C_B shunts a large proportion of input signal power into R_S where it is dissipated. For a large local-oscillator drive R_B becomes smaller, the shunting effect of C_B is reduced, but in the limit the value of R_B approaches that of R_S and the power is shared.

Combining these effects it is obvious that for any values of R_B, C_B, and R_S there is an optimum local-oscillator drive for minimum conversion loss and that minimization of the product $C_B R_S$ leads to the lowest conversion loss.

The ratio of power available to the nonlinear barrier resistance P_B to the total power available P can be calculated.

$$\frac{P_B}{P} = \left(1 + \frac{R_S}{R_B} + \omega^2 C_B^2 R_B R_S\right)^{-1}$$

where $\omega = 2\pi f$.

The local-oscillator power and the bias applied to the diode are adjusted to optimize R_B. The local-oscillator drive should be sufficiently large to provide the necessary degree of nonlinearity but not so large that R_B becomes comparable or smaller than R_S.

2.1.2 *Noise*

When excited by a direct or alternating voltage, a mixer crystal produces more noise

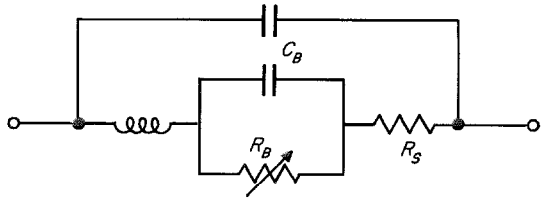


Figure 2—Equivalent circuit of a crystal diode including capacitance of case and inductance of connections and whisker.

than an equivalent resistor at the same temperature. This noise has three main sources: (A) thermal or Johnson noise, (B) shot noise, and (C) flicker noise. The last is small enough at high frequencies to be neglected. Thermal noise is generated mainly in the total series resistance. Uhlir [1] has postulated that shot noise is associated with minority-carrier flow across the barrier and may happen in three main ways: (A) a thermally generated minority carrier may diffuse to the barrier where it falls down the potential hill, (B) a majority carrier may possess sufficient energy to overcome the potential barrier to become a minority carrier, and (C) this minority carrier, if it does not recombine, may recross the barrier. The total barrier noise is the sum of these three components, the origins being principally the emission admittances of both types of carrier and the flow of saturation current. To reduce, as much as possible, the incidence of recrossing the barrier by minority carriers, the carrier lifetime in the semiconductor material should be as short as possible. It can be shown that both Johnson noise and shot noise can be minimized by holding R_S at a low level and that small saturation current or a good reverse characteristic is consistent with low shot noise. It is obvious that a reduction of Johnson noise will follow a reduced R_S but the bearing of this factor on the shot-noise power is not so obvious. Uhlir showed that for an ideal nonlinear resistance (barrier only) lowest shot noise occurs at infinite local-oscillator swing. This, however, is modified by the presence of series resistance. An increase in local-oscillator swing results in a reduced R_B, an increase in shot-noise current, but a decreased R_B shot-noise power contribution. When R_B falls to a value comparable with that of R_S, the R_S and R_B contributions become comparable and because the noise power contribution of R_S increases with shot-noise current while that of R_B decreases, a situation is reached where further increase of local-oscillator drive results in increased noise. As in the case of conversion loss, therefore, there is an optimum value of local-oscillator drive

for minimum noise and it is fortunate that the two requirements are not incompatible. The importance of a low value of series resistance is again emphasized. Shot noise associated with high reverse saturation current has been noted in practice. Lindell and Oxley [2] have reported that when a crystal diode is alternately and equally biased forward and backward the noise output for current in the reverse direction is considerably greater than for current in the forward direction, although the reverse current is very much smaller. Further, the greater the reverse current the greater is the noise produced.

2.1.3 *Radio-Frequency Impedance*

In considering the radio-frequency impedance of a crystal mixer the equivalent circuit of Figure 1 must be modified to include the overall case capacitance and the inductances of the whisker and internal connections. These are included in Figure 2. These additional factors can be tuned out in the crystal-holder design but have to be held constant from sample to sample since it is the spread in radio-frequency impedance in a given design of holder that is of significance. In fact, a consistent radio-frequency impedance is chiefly a matter of precision in manufacture. This precision applies not only to the mechanical tolerances of component parts, but accurate assembly to a pattern, control of semiconductor properties, surface treatment, whisker-point shape, whisker pressure, and forming.

2.1.4 *Intermediate-Frequency Impedance*

The reactive components considered in 2.1.3 are too small to be very significant at 70 megacycles per second, the intermediate frequency, and the same is true of the effects of the radio-frequency system in the low-Q case as implied in the present broad-band application. Thus the intermediate-frequency impedance is almost wholly resistive and corresponds closely to the average resistance when the characteristic is traversed by an alternating voltage. R_S, being only a few ohms, has a negligible effect so that the intermediate-frequency impedance is clearly dependent on R_B the barrier resistance.

It will be remembered that the value of R_B decreases with increasing local-oscillator drive and this must also be true of the intermediate-frequency impedance. It is necessary, therefore, to arrange in design that there is correspondence between optimum conversion loss, minimum noise, and the required intermediate-frequency impedance at the chosen local-oscillator drive level. Junction area obviously also bears on intermediate-frequency impedance but there is a very distinct limit to the permissible increase in C_B resulting from any effort to reduce the intermediate-frequency impedance by an increase of junction area.

2.1.5 *Burnout Level*

Burnout is a condition due to overload resulting in a significant deterioration of mixer performance. It has become customary in radar usage to specify the maximum spike energy (measured in ergs) to which the crystal diode may be submitted. This is because the radar crystal is subjected to such spikes due to leakage past the transmit-receive cell at each transmitted pulse. This kind of rating is not relevant to crystals used in communication systems, a continuous-wave overload rating being more appropriate.

Burnout is entirely a function of heat dissipation at the junction and depends on the semiconductor material, the area of the junction, and whether or not the junction was subjected to prior electrical forming.

2.2 Practical Mixer

The structure and mode of operation of the point-contact microwave mixer diode are not, even now, unambiguously understood so that, except for the guiding principles outlined in the previous section and the experience of

other workers, the ultimate approach is empirical to an appreciable extent. It is evident that a small $R_S C_B$ product and particularly a small R_S should be significantly important.

2.2.1 *Choice of Semiconductor*

Messenger and McCoy [3] in an excellent treatise on microwave mixer diodes have shown that maximization of the quantity $n^{1/2}b$, where n is the impurity concentration and b the carrier mobility, is desirable. The quantity bears strongly on the bulk series resistance R_S which, as has been shown, must be minimized.

Microwave mixers in the past have been made from p-type silicon and n-type germanium. Experimental units have been made from n-type gallium arsenide. From considerations of mobility the choice would be n-type since electron mobility is always higher than hole mobility. Silicon of n-type, however, turns out to be poor for reasons not well understood. Electron mobilities in germanium and gallium arsenide are quoted as 3800 and 5000 square centimetres per volt-second, respectively, while hole mobility in silicon is 500 square centimetres per volt-second. This is why series resistance in silicon microwave diodes ranges from 15 to 75 ohms compared with 3 to 5 ohms for germanium and gallium arsenide. On this evidence it would appear that, in this respect at least, silicon is inferior. On the question of intermediate-frequency impedance too, a special problem in our case, silicon diodes are at a disadvantage, showing higher values than those of germanium and gallium arsenide. The popularity of the silicon diode can perhaps be explained by its simplicity in manufacture consisting, as it does, of a tungsten whisker merely "shaken down" on to a die of polycrystalline silicon. In contrast to this, germanium and gallium arsenide are invariably monocrystalline and the whisker has to be welded or formed by the passage of an electric current to the semiconductor surface. A mechanical advantage of this forming is that it confers a degree of ruggedness not present in the silicon diode.

Other factors for comparison are lifetime, energy gap, and dielectric constant. Lifetime should be short to avoid charge storage and transit-time effects, energy gap should be large to ensure high working temperature, and the dielectric constant should be small to minimize barrier capacitance. At the high impurity concentration levels involved lifetime is so short, due to the density of recombination centres, that it is difficult to measure.

On energy gap and dielectric constant gallium arsenide is seen from Table 1 to score over both germanium and silicon.

Reasons for the rejection of silicon have been given but the fact that germanium was chosen in face of the apparent superiority of gallium arsenide requires justification. The reasons are as follows.

(**A**) The much-greater background of metallurgical experience with germanium.

(**B**) The apparent inability to achieve the promised electron-mobility advantage in heavily doped gallium arsenide.

(**C**) The very-marginal superiority of gallium-arsenide diodes reported by other workers and then only in selected laboratory samples.

TABLE 1
PROPERTIES OF SEMICONDUCTORS

Material	Energy Gap in Electron-Volts	Dielectric Constant
Germanium	0·68	16·0
Silicon	1·1	11·8
Gallium Arsenide	1·35	11·1

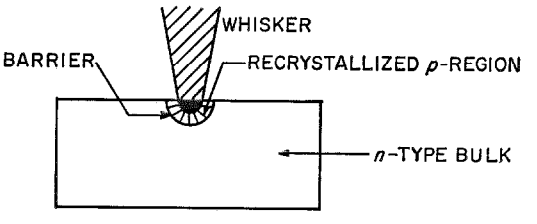


Figure 3—Model of a diode.

(**D**) The report that one reputable American manufacturer had withdrawn from the market its gallium-arsenide microwave diodes.

(**E**) Considerable doubts about the long-term stability of other heavily doped abrupt-junction gallium-arsenide devices.

(**F**) Our own inability to achieve results superior to germanium in our early investigations coupled with the doubts above.

2.2.2 *Whisker Material*

The function of the cat whisker in the formed germanium diode differs from that in the unformed silicon diode. The latter is an example of a metal-to-semiconductor contact diode while in the former, the semiconductor immediately under the whisker undergoes a change due to the heat developed by the passage of the forming current. In the extreme case (heavy forming), the conductivity type is changed and a *p-n* junction results. In the case of the microwave diode where the forming has to be extremely light to preserve a small low-capacitance junction and where the doping level of the *n*-type bulk semiconductor is high, it is doubtful whether complete conversion takes place. A diode still results but of the contact type with probably a region of higher-resistivity *n*-type material immediately under the point. This is not a desirable state of affairs since this region of higher resistivity adds to the unwanted series resistance R_S. The conversion to *p*-type can, however, be enhanced by the inclusion in the whisker material of a metal of Group III in the periodic table, for example indium, gallium, or aluminium, which are metals commonly used as *p*-type doping agents in germanium. When such a metal is included in sufficient quantity a minute *p*-type region of low resistivity results and an abrupt *p-n* junction is formed. The increased series resistance due to alloying an undoped whisker is largely avoided.

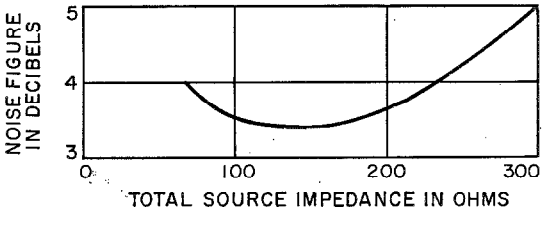


Figure 4—Noise figure at an intermediate frequency of 70 megacycles per second.

Mechanically a whisker wire has to be resilient, springy, and capable of supporting a sharp point under some pressure. A wire having the requisite electrical and mechanical properties is composed of 90-per-cent copper and 10-per-cent aluminium. Such a composition is known as aluminium-bronze and forms the whisker wire in this diode.

A model of this diode is depicted in Figure 3. It is assumed that a convergent flow of electrons takes place from the highly doped *n*-type region to the less-heavily doped *p*-region. It is probable that lifetime is very short in the aluminium-copper-doped *p*-region and recombination can be speeded by providing a germanium surface of high recombination velocity, that is, mechanically polished.

2.2.3 *Intermediate-Frequency-Impedance Realization*

The necessity for matching the intermediate-frequency terminals of the mixer to the intermediate-frequency amplifier is of prime importance. Modern grounded-grid low-input-impedance amplifiers demand either mixers of correspondingly low intermediate-frequency impedance or matching transformers or networks. Such matching transformers are undesirable in broad-band working, hence the request for low intermediate-frequency impedance in the mixer. The value of 100 to 150 ohms is the optimum for minimum noise in the present case as is seen in Figure 4.

It was shown in Section 2.1.4 that, apart from effects of the radio-frequency circuit which are fixed by other considerations, the intermediate-frequency impedance is largely

dependent on the barrier resistance at the working local-oscillator drive. The barrier resistance also depends on the resistivity of the semiconductor and the junction area or, as is equivalent, on the forward conductance. Most germanium microwave mixers have intermediate-frequency impedances of about 200 ohms and this was confirmed in our early work. Occasionally a diode of 150 ohms would appear but generally speaking the trend of the curves of Figure 5 was confirmed. These curves show rectification efficiency and intermediate-frequency impedance against forming current for 1 milliwatt drive at 6 gigacycles per second (increased forming current means larger junction area). The intermediate-frequency impedance is seen to continue to fall with increased forming but rectification efficiency (and presumably smallest conversion loss and optimum junction area) passes through a well-defined maximum.

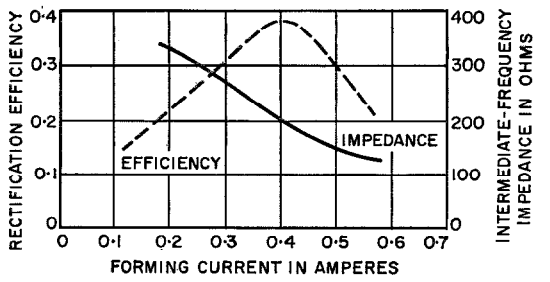


Figure 5—Rectification efficiency and intermediate-frequency impedance plotted against forming current for a single-junction germanium diode for 1 milliwatt drive at 6 gigacycles per second.

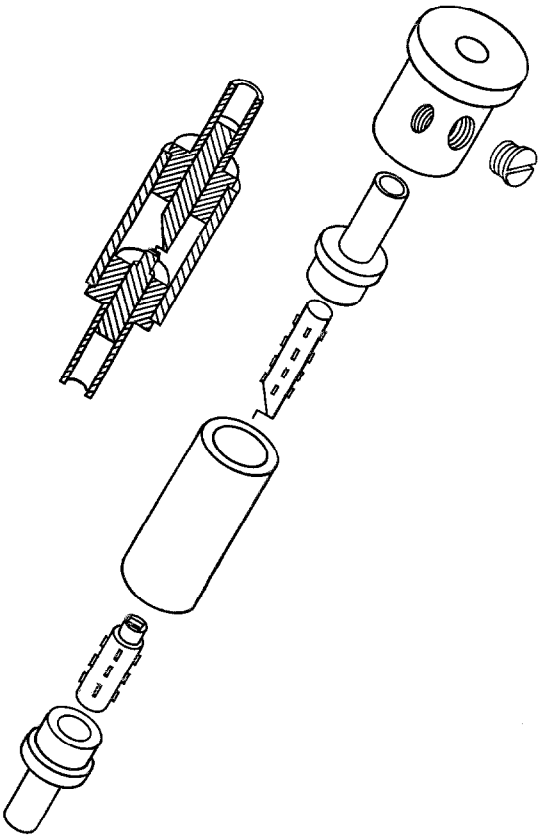


Figure 6—Final form of diode.

It should be noted that to keep R_S low the germanium resistivity was as low as would support efficient rectification.

The reason for the degradation of conversion loss with increase of junction area can be seen from the following relations, which apply to a diode of this geometrical form:—

$$R_S \propto 1/a$$
$$C_B \propto a^2$$

where a is the radius of the junction.

The disproportionate change in R_S and C_B for an increase in junction area means a higher R_SC_B product and an increased conversion loss.

If now the effective junction area is, instead, increased by putting down two points on to the germanium surface, the changes in R_S and C_B are proportionate, the product R_SC_B is unchanged but the impedance is halved.

2.2.4 *Final Diode*

Figure 6 shows the final form of the diode together with its component parts. The semiconductor die is of n-type germanium, having a resistivity of 0·004 ohm-centimeter and the twin cat whisker is of aluminium-bronze (90-per-cent copper and 10-per-cent aluminium). The cartridge is of the standard form adopted

by most diode manufacturers, is symmetrical, and has the facility of polarity reversal by virtue of a separate end cap.

Tests at microwave frequencies are complicated and would warrant a separate report. Figures 7, 8, and 9 show typical results for conversion loss, over-all noise figure, and intermediate-frequency impedance respectively. These are for an average diode; conversion losses of 3·9 decibels have been measured.

The curves show that the target specifications for conversion loss, over-all noise figure, and intermediate-frequency impedance have been met for a local-oscillator drive of 1·2 milliwatts and a direct-current forward bias of 0·1 volt. The actual over-all noise figure is degraded by the poor noise figure of 3·3 decibels of the amplifier used for the tests.

Substitution of the figures for conversion loss and over-all noise figure in the equation in Section 2.1 will reveal a rather larger value for the noise-temperature ratio t_x than that specified. The overriding effect of conversion loss has to be taken into account in this connection and the possibility of a slightly optimistic conversion-loss figure has to be admitted. No independent measurement of t_x has been made.

The double-whisker structure ensures a burn-out figure much better than the specification.

3. Medium-Level Up-Converter

It is recognized that up-converters using modern non-linear-capacitance diodes are more efficient than the older converters employing non-linear-resistance action. It is pertinent, however, to weigh this disadvantage against certain other desirable effects that can more-easily be realized in the non-linear resistance. The destructive influence of the parasitic bulk series resistance noted in the down-converter can be used to good effect in the up-converter to bring about an early saturation condition, which is useful for its amplitude-limiting action. Degrading the rectifying action at the higher frequencies results in lower levels of interfering harmonic production.

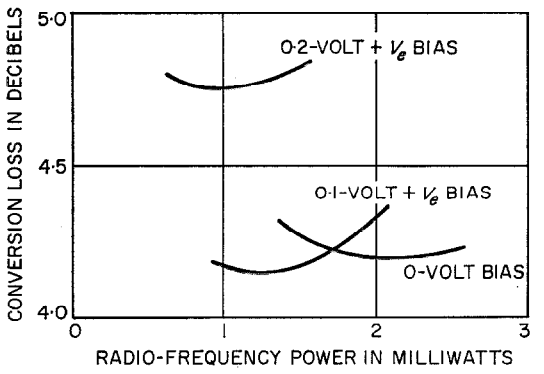


Figure 7—Conversion loss versus radio-frequency power input as a function of bias.

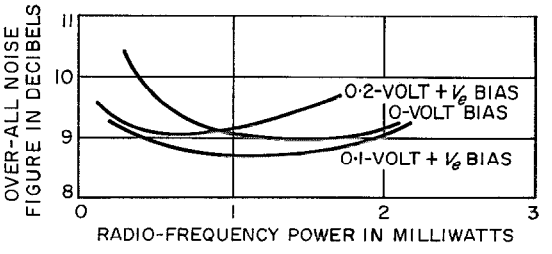


Figure 8—Over-all noise figure plotted against input power. The amplifier had a noise figure of 3·3 decibels. Three bias values are shown.

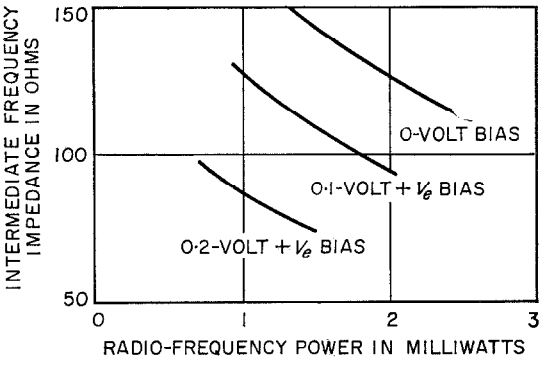
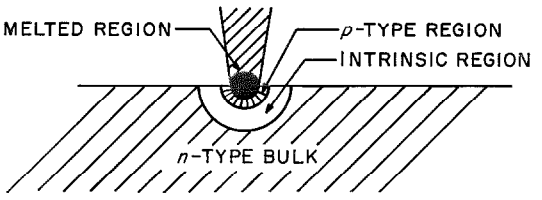


Figure 9—Intermediate-frequency impedance as a function of input power for the indicated bias values.

Figure 10—Formed point-contact germanium diode.



3.1 Design Parameters

The up-converter may better be regarded as a modulator in which the impedance of the diode to the high frequency is varied by the incident intermediate-frequency signal. If the diode resistance could be swung from open circuit to short circuit then the conversion could be made nearly lossless. It is important to note here that we are considering the loss from in to out at super-high frequencies, that is, from unmodulated carrier in to frequency-modulated carrier out.

In a typical circuit arrangement a hybrid T is used as a balanced mixer, the diodes are mounted in the side arms at points equidistant from the centre of the junction and are energized in antiphase at the intermediate frequency. The carrier is introduced in the H arm and the wanted sideband is taken out of the E arm through a filter.

3.1.1 *Conversion Loss*

The model of Figure 1 is equally applicable to the up-converter. If pure-resistance action is desired it is evident that the change in R_B must be maximized while C_B should preferably be as invariant as possible. This means that both the forward conductance and the reverse resistance should be as high as possible and that these extreme conditions should be achieved at the required intermediate-frequency swing. The junction capacitance can be larger than in the down-converter but should not be so large as to shunt seriously the barrier resistance in the reverse-biased condition.

3.1.2 *Limiting*

A saturation effect occurs when the intermediate-frequency swing encroaches on the linear region of the forward characteristic. As in the down-converter R_B reduces with increase of drive until a point is reached where R_S takes over and further increases in drive level produce little effect on the impedance. The incidence of this saturation can be controlled to a large degree by choice of semiconductor resistivity, and hence R_B, and by designing deliberately for a high value of R_S. In this way the rectifying property of the diode at high frequencies is also degraded.

3.2 Practical Up-Converter

The *p-i-n* structure, in which an intrinsic layer separates the p and n regions, has all the attributes for a variable-resistance modulator.

The intrinsic layer ensures a high reverse resistance and permits a larger junction area for a given capacitance. The capacitance, too, is not so voltage dependent and is largely invariant. During the forward half-cycles of the intermediate-frequency drive the intrinsic region is flooded with carriers and its conductance is high. At the super-high frequencies transit time across the intrinsic region precludes efficient rectification.

3.2.1 *Realization of p-i-n Structure*

In the formed point-contact germanium diode a variety of effects can be obtained by choice of the parameters, starting type, resistivity, whisker composition, and degree of forming. It is well known that when n-type germanium is heated sufficiently the conductivity decreases and, depending on the temperature to which it is raised, eventually goes through an intrinsic state to the opposite p conductivity type (thermal conversion). By correct choice of the parameters mentioned the structure of Figure 10 can be realized. For this case a comparatively high-resistivity n-type germanium is chosen in combination with a whisker carrying a p-type impurity metal. The degree

of forming is chosen to produce a heavily doped p region immediately under the whisker tapering into a region where the conductivity types are balanced and therefore of high resistivity. The transition can be made relatively sharp if a square pulse of forming current is applied having a rapid fall time. In this model, as distinct from that of the down converter, the relative resistivities of the p and n regions would suggest a divergent flow of holes from the p region into the bulk.

3.2.2 *Results*

From the Smith chart of Figure 11 the impedance change from +1 volt to −1·6 volts is seen to be predominantly resistive as expected. The graph showing conversion loss against intermediate-frequency drive, Figure 12, demonstrates that the conversion loss of 7·5 decibels at the required saturation level is better than the specification by 3·5 decibels. The required compression of 5 : 1 is maintained for an input power change of 5·6 decibels which is considered adequate.

No interfering harmonics of the intermediate frequency were detected at −117 decibels, the limit of the measuring equipment.

4. Acknowledgements

The author wishes to recognize the valuable assistance of A. B. Kaiser in making the diodes, Mrs. E. B. Evans and G. M. Meathrel

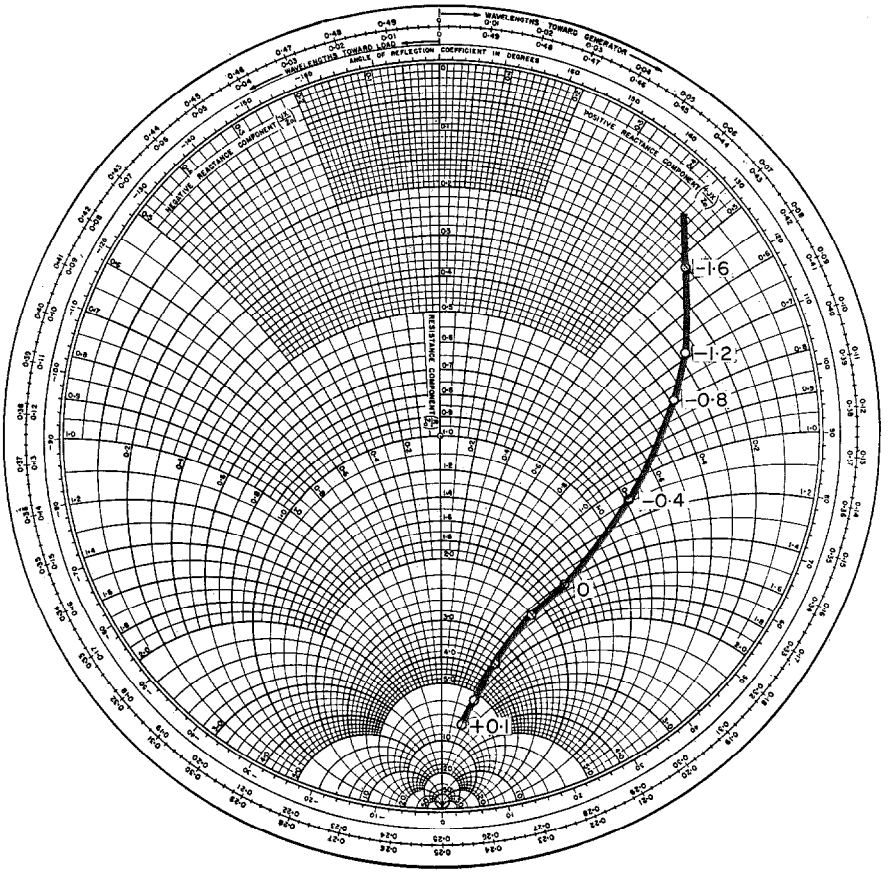


Figure 11—Smith chart plotting impedance change as a function of bias.

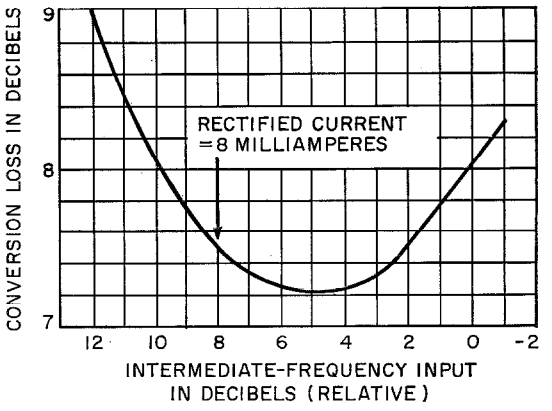


Figure 12—Conversion loss as a function of intermediate-frequency drive at 70 megacycles per second with 200 ohms of series resistance and a super-high-frequency output at +15 decibels referred to 1 milliwatt.

for work on whiskers, B. F. Armsby for making the tools, and D. R. Hill and R. C. Baron for painstaking measurements.

5. References

1. A. Uhlir, "Shot Noise in *p-n* Junction Frequency Converters," *Bell System Technical Journal*, volume 37, pages 951–988; July 1958.

2. A. Lindell and T. Oxley, "Radar Crystal Valves," *Proceedings of The Institution of Electrical Engineers*, Part B, Supplement 15, volume 106, pages 426–433; May 1959.

3. G. C. Messenger and C. T. McCoy, "Theory and Operation of Crystal Diodes as Mixers," *Proceedings of the IRE*, volume 45, pages 1269–1283; September 1957.

J. B. Setchfield was born in Nottingham, England, on 9 February 1907. Starting as an amateur, he joined the radio industry in 1924 and until 1940 served with several firms in the development of broadcast receivers.

From 1940 to 1945, he was in the Royal Air Force, during the last half of which he was at the Telecommunication Research Establishment developing airborne radar equipment.

He joined Standard Telecommunication Laboratories late in 1945 and now heads a group studying the application of solid-state devices to microwave systems.

Vacuum Tubes for Submerged Repeaters

F. G. HAEGELE
Standard Telephones and Cables Limited; London, England

The cost of an interruption in service and of replacing a submerged repeater in a multichannel telephone submarine-cable system places fantastic importance on the reliability of all parts of such repeaters and particularly on the vacuum tubes that are the essential active elements of these installations.

1. Properties of Submarine Repeater Tubes

1.1 Electrical

Unusual conditions must be anticipated in the design of vacuum tubes for long submarine-cable communication systems as they will be contained in amplifiers in sealed vessels at the bottom of the ocean accessible electrically only over the coaxial cable that connects the repeaters to each other and ultimately to the shore-based terminal equipment. Consequently, all power must be derived from a voltage drop across each repeater produced by current passing through the cable.

Calculations, as applied to optimized amplifiers, indicate need for high transconductance, low interelectrode capacitances, and low electrode voltages and currents.

Other things being equal, the higher the gain of the tube, the fewer will be the number of repeaters in the cable, provided that, at any point in the system, the signal-to-noise ratio does not fall below the specified limit. This calls for high transconductance. The lower the interelectrode capacitances, relative to high transconductance, the greater the transmitted bandwidth.

As all the amplifiers are operated in series from a voltage applied across the entire cable, low electrode voltages minimize this over-all voltage and low currents, particularly heater current, further reduce the operating power.

These basic properties must be associated with the two features, long life and utmost reliability, that put submarine repeater tubes in a class of their own.

1.2 Long Life

Long life implies maintenance of thermionic emission, constancy of electrical parameters, maintenance of insulation, and low evaporation rate from hot components.

1.3 Reliability

Reliability implies the possibility of an accident, a failure that is sudden and unexpected. Nevertheless a probability of failure can be assigned to a tube based on tests of a quantity of them.

2. Fundamentals of Design

2.1 General

The requisite properties of the tube are inextricably linked in a compromise, the success of which depends on the designer's skill in assessing the merits of conflicting factors. They will be treated generally here. Specific examples will appear in the later sections.

2.2 Electrical Characteristics and Dimensions

A high transconductance demands above all a close control of grid-to-cathode spacing. Small spacing presents insulation risks particularly in the presence of accidentally introduced conducting particles. If the spacing is less than the pitch of the winding of the grid wire non-linear effects result from the inhomogeneous electric field. Thin grid wires permit closer winding pitch but introduce mechanical weakness.

Low capacitances demand that only the active parts of relevant electrodes be exposed to one another. In practice, of course, connecting leads and essential screens make a substantial contribution to these capacitances.

Low electrode voltages demand closer electrode spacings to maintain the requisite electric field strengths, thus adding to capacitance.

Low heater currents are limited both by power

requirements and by the mechanical weakness of thin wire and evaporation if it is too thin.

Generally low-voltage operation increases the problem of maintaining uniformity among individual tubes as well as exposing the structure to greater risk of failure from accidentally introduced particulate contamination.

2.3 Components for Long Life

The outstanding requirement for long life is the maintenance of an adequate emission from the cathode.

Foremost, the alkaline earth carbonates must be satisfactory for this purpose by virtue of high purity and low evaporation rate. Secondly, the metal of the cathode core must be suitable. A vast amount of experience has been, and is still being, amassed on the constitution of the cathode core metal and its intentional and unintentional admixtures. Indeed, as the following history of submarine repeater tubes will reveal, certain aspects of the behaviour of cathodes were not clearly appreciated at the time and circumventing and limiting means were employed to avoid the deleterious phenomena observed.

For long life, the cathode must operate at the lowest possible temperature that will maintain stable equilibrium with the deactivating effect of residual gases in the tube. Evaporation is then least. In practice, with well-processed high-quality materials, the chosen optimum is 750 degrees centigrade.

Although not fundamentally an emissive property of the cathode, it is known that constituents of the cathode core metal can react, during operation, with the oxide coating to produce an interface layer having properties of an ohmic resistance. This causes degeneration in the cathode circuit and consequent loss of gain in the amplifier stage. Several new alloys have been devised since 1950 to avoid this effect.

A similar degeneration is brought about by the intrinsic resistance of the oxide layer. Control of this is partly one of correct coating and partly one of avoidance of gas contamination, which de-activates the layer and thus raises its resistance.

Aside from the deterioration problems of the cathode, the heater also has a wear-out function. Experience dictates the use of the lowest heater temperature that will maintain correct cathode operation.

A final point in protracted maintenance of constant tube characteristics is the possibility of relaxation of accidentally stressed components by the heat of operation producing an abnormal trend of characteristics during life.

2.4 Factors Affecting Reliability

The concept of reliability as the inverse of accident leads to the view that it can be obtained by perfection of all aspects of the tube and its preparation. Given the optimum design, it is necessary that the materials be as specified and that the execution of all the processes be likewise above criticism.

Translated into statistical terms, if the probability of failure of a tube be p then, in a system containing n repeaters each carrying redundancy in the form of twin parallel 3-tube amplifiers, the probability of breakdown of the system is $9\,np^2$. (A simple derivation of this expression is given in Section 9.) Thus, in a system containing 100 repeaters, the probability of the system failing is approximately equal to the probability of failure of an individual tube when the latter is 0·999 reliable over the acceptable operating life of the system.

It can be deduced that to attain a reliability at this level, the conventional system of statistical sampling becomes meaningless. The sample to be tested would be so large that mass production would be needed to handle it and the object would be defeated by the economics, at least, of controlling the operation at that level. Alternatively, accepting that, since accidents are random and their causes manifold, with a fairly large number of tubes the accident probability density will be uniform through

time, it might be concluded that the solution lies in this direction. Calculation again gives an impracticable answer. The time to life-test a manageable sample to give a reasonable certainty would exceed the acceptable life of the system (which is, say 20 years).

From these considerations one is driven inexorably to the conclusion that the desired reliability can be attained only by perfection in manufacture by all possible means and to seek, by virtue of the uniformity gained, to control the quality by supervision and elimination of disturbing influences. The practical realization of this is shown in Section 4 on manufacture.

3. Historical Survey

3.1 General Remarks

The physical form of the tubes concerned in this survey is shown in Figure 1. Table 1 gives details of their characteristics. The following notes on the individual types trace evolution to the present day and draw attention to the effect of fundamental discoveries on cathode behaviour and the advent of new constructional techniques.

3.2 *5A/153* Pentode

The *5A/153* pentode was a direct outcome of experience on coaxial repeater pentodes for land-based repeaters. At this time the basic known concepts for long life were low-current-density operation, low power dissipation, and low collector-electrode voltages. The phenomenon of interface growth was not recognized but selected batches of cathode nickel were known by experience to give consistent emission and small transconductance fall during running as well as low evaporation rates.

This type was used in 3-tube repeaters that served successfully for 10 years in 36-channel shallow-water cables. At the end of this time the repeaters were replaced by more-modern examples containing *5A/162* tubes to extend the capacity to 180 channels.

3.3 *5A/161* Pentode

As can be seen from the photograph the *5A/161* was generally similar to the *5A/153* type. The principal differences were small modifications to the grids to suit operation at a lower voltage and lower anode current. The heater was also altered to adapt it to the voltage needs of the repeater.

TABLE 1
Characteristics of Submarine-Repeater Tubes

	Type and Year Introduced						
	5A/153 1947	*5A/161* 1949	*5A/162* 1952	*5A/181* 1958	*5A/182* 1958	*LS882* 1958	*VX7158* 1960
Heater Voltage in Volts	10	25	5·5	10	10	10	7·4
Heater Current in Milliamperes	430	180	265	300	300	300	390
Anode Voltage in Volts	150	75	90	70	40	70	50
Screen Voltage in Volts	150	75	60	70	50	70	45
Anode Current in Milliamperes	9	2	6	10	3	10	15
Screen Current in Milliamperes	2·5	0·5	1·6	2·5	0·7	2·2	3
Transconductance in Milliamperes Per Volt	7·5	3·5	6	6·9	5·9	17	22
Control-Grid Voltage in Volts	−3	−2·5	−1·6	−3·5	−1·5	−1·7	−1·8
Anode Impedance in Kilohms	500	800	300	150	250	200	45
Input Capacitance in Picofarads	13·7	13·7	7·5	9·8	10·5	16	24·5
Output Capacitance in Picofarads	8	8	5	6·2	7·6	3·5	5
Anode-Grid Capacitance in Picofarads	0·007	0·007	0·020	0·020	0·025	0·020	0·025

The application of the repeaters was unusual. A 3-stage push-pull amplifier (6 tubes) was located at the continental-shelf edge at either end of a transatlantic telegraph cable. The greatly attenuated signal from the deep-sea portion of the cable was boosted and shaped before passing along the shallow-water section where ships produce electric interference. The resultant improvement in signal-to-noise ratio permitted the speed of the cable to be increased.

3.4 *5A/162* Pentode

The *5A/162* is equivalent to the British Post Office *6P12* tube [1]. Fundamentally, the electrical characteristics are adapted from the *6AM6* (*CV138*) miniature high-transconductance pentode to operate at the low anode and screen voltages indicated. Two noteworthy features of this tube made it long lived. A platinum cathode sleeve was used to avoid interface growth and very thorough attention was paid to elimination of gas generation [2], deleterious to cathode emission, by material purity and processing.

3.5 *5A/181* and *5A/182* Pentodes

These tubes are the respective equivalents of the British Post Office types *10P1* and *10P2* tubes [3]. They are respectively output- and input-stage versions of a common mechanical construction and are direct descendants of the *5A/162*. Further attention was applied to poisoning influences on the cathode, which is of the same tungsten-nickel alloy as a later modification of the *5A/162*. The heater-cathode insulation has been vastly improved by use of an intervening non-porous alumina sleeve to eliminate any danger of insulation breakdown inherent in operation with heater positive to cathode. This permits development of the anode voltage across the heater chain instead of in series as previously necessary.

3.6 *LS882* Pentode

In 1956 consideration was being given to the possibility of increasing the channel capacity

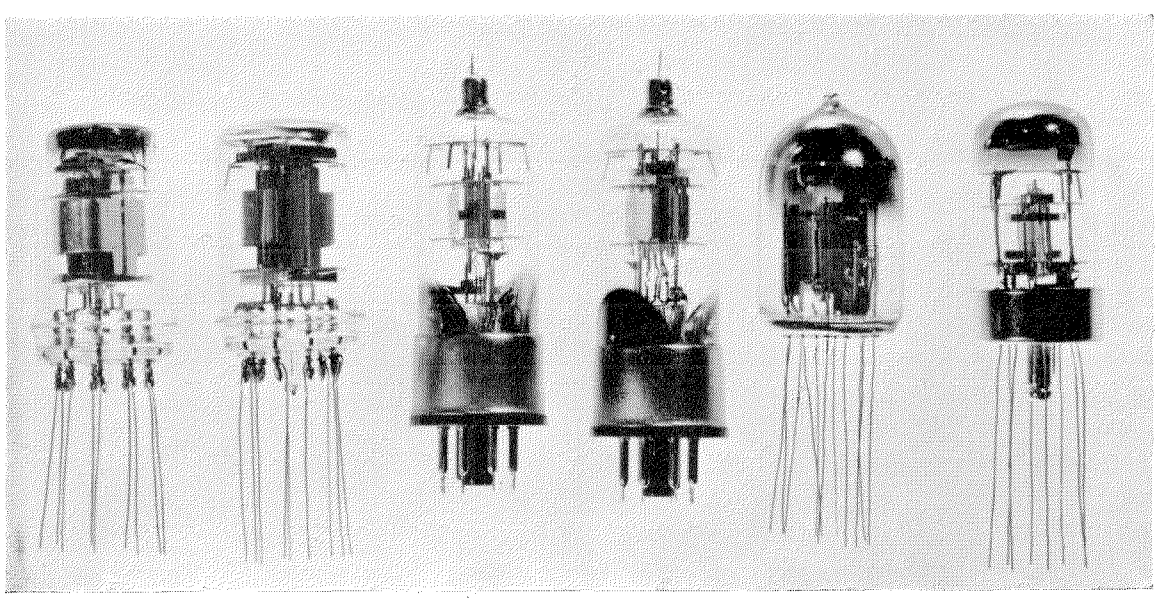


Figure 1—These submarine-repeater tubes are, from left to right, *5A/153*, *5A/161*, *5A/162*, *5A/181–5A/182*, *LS882*, and *VX7158*.

of submarine cables beyond that feasible with existing tubes. The characteristics of the *5A/162* and *5A/181* class of tube limit bandwidth to 0·5 megacycle per second for repeater spacing of 20 nautical miles. (A wider frequency band is possible by reducing the spacing since the cable attenuation is proportional to the square root of frequency.)

It was clearly recognized that no substantial improvement could be obtained with existing designs of tubes and that a parallel development in conventional repeater tubes must be exploited. The frame-grid construction [4, 5] offered, as it had for land-based repeaters, a large increase in transconductance with relatively little change in the important capacitances.

Calculations showed that, at the low working voltage, the spacing between control grid and screen grid would be quite small and the use of a conventional self-supporting screen grid would introduce electric field inhomogeneity and consequent lower efficiency. It was decided, therefore, to employ a frame grid for this grid also.

The resulting characteristics show the tube to be suitable for an amplifier having a bandwidth of 2 megacycles per second and providing 200 to 240 channels. This tube was the prototype of the *VX7158*.

3.7 *VX7158* Tetrode

The development exercise on the *LS882* proved the feasibility of the new construction. However, it appeared that the working voltage of the *LS882* would make the over-all voltage on long-haul systems excessive. Consideration was given to applying the principles of the *LS882* to a lower-voltage tube. A theoretical and practical investigation showed that the desired characteristics could be best obtained with a tetrode structure.

It may be remarked that in tubes designed for greater bandwidths, such as the *LS882* and the *VX7158*, the phase angle must be held to a low value since low-voltage operation gives a substantial transit angle. For this reason inductance of the leads must be minimized and a pressed glass base is indispensible.

4. Manufacture

4.1 General Remarks

It will be clear from the foregoing discussion that successful manufacture of submarine-repeater tubes calls for the highest standards in cleanliness of components and of fabrication. Coupled with this is the closest attention to dimensional control to ensure uniformity of the product.

The underlying principles involved in the design of the manufacturing section were: complete separation from other factory activities; restricted access by location and control of entry; ordered arrangement of processes; isolation of the assembly area, which has its own cleaned-air system; provision for close inspection, under engineering control, of all stages of manufacture; and adjacent testing areas, including life test, for rapid feedback of results.

4.2 Procedure

The layout of the section is shown in Figure 2. Entrance to the area is restricted to properly garbed personnel by a steward stationed in the store immediately adjacent to the double entry doors. The operation is arranged in a logical sequence so that there is an ordered flow of the components towards the assembly area, which is pressurized so that air flow into the rest of the system is counter to the component flow. (The latter is indicated by the thick line in the figure.)

Practically all the piece parts are fabricated from the raw material within the piece-part section on hand presses or winding machines as appropriate. This ensures that there is no danger of "cross-infection" from factory machinery used with other materials and processes.

The inherent relevant impurity content of

piece parts (mostly gas such as oxygen and carbon monoxide) must be relentlessly eliminated. The unwanted gas content is reduced to less than 2 percent of the original amount. This compares with 90 percent in ordinary receiving tubes.

The operators enter the assembly area through air locks in which they change their nylon coats for fresh more-thoroughly enveloping apparel including shoes and headgear.

Inspectors in the room check each assembler's work, providing immediate information on quality.

The subsequent pumping operation is lengthy and designed to expel the last dregs of gas from the tube.

4.3 Testing

All electrical tests are performed on high-precision equipment capable of measuring transconductance to 4 significant figures. A second mechanical inspection is made prior to life testing.

Life testing takes place on racks with closely controlled current supplies, where batches are run for a period of approximately a year. The tubes are tested at intervals during this time to gauge the trends of the principal parameters. Finally, on the basis of satisfactory trends, tubes are selected for repeater use. Tubes from the same batches are retained as shadow life tests to provide background information on the behaviour of their submerged counterparts. The final selection takes into account not only the electrical trends observed during the year but also the mechanical inspection results, which are repeated and counterchecked at the end of the period. The final tests are supplemented by X-ray stereo photographs to reveal hidden components, such as heaters and lead wires in the base and by special electrical tests, as for harmonic generation and noise, dictated by the specific circuit requirements.

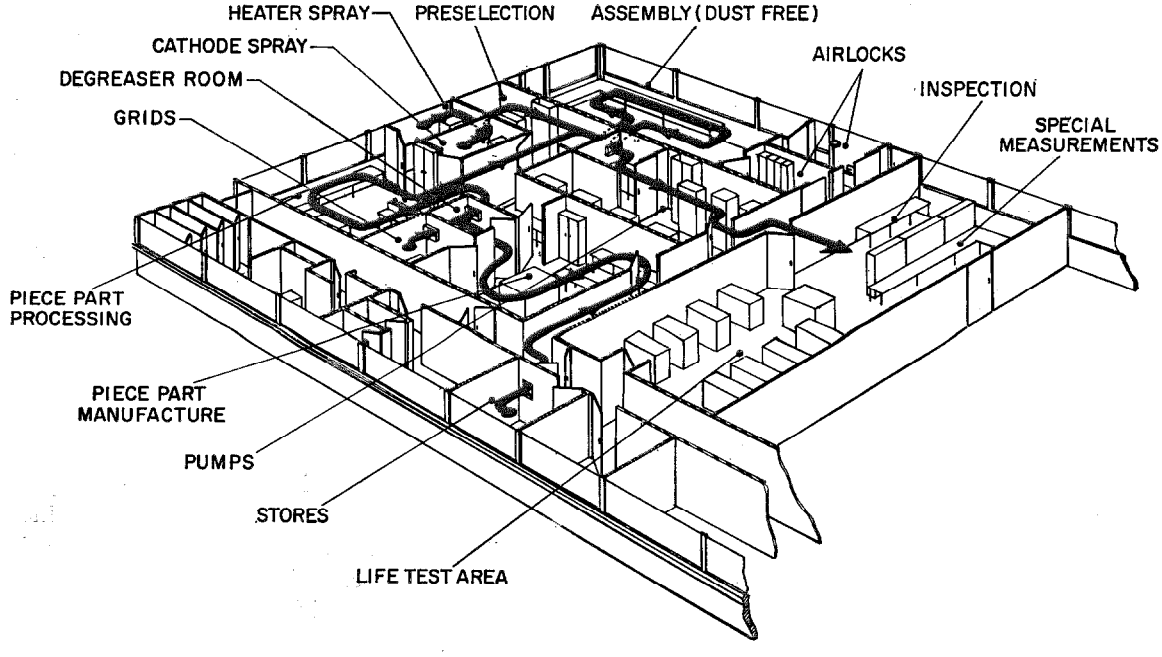


Figure 2—Layout of manufacturing area.

5. Performance

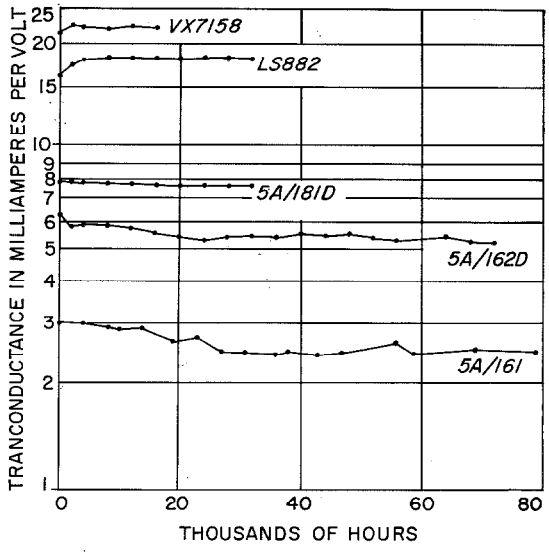
5.1 Uniformity of Product

In view of the need to maintain the tube parameters within close limits despite the complication introduced by low-voltage operation, it is worth demonstrating by a histogram what has been achieved. Figure 3 shows the control exerted over transconductance.

5.2 Life Performance

The shadow life tests on the various types continue to provide impressive evidence on the long-term stability of the tubes described, as Figure 4 illustrates. As it is difficult to present the very-small changes adequately it is worth remarking that a batch life test on *5A/181* tubes of early manufacture, which has now

Figure 4—Transconductance trends with life.



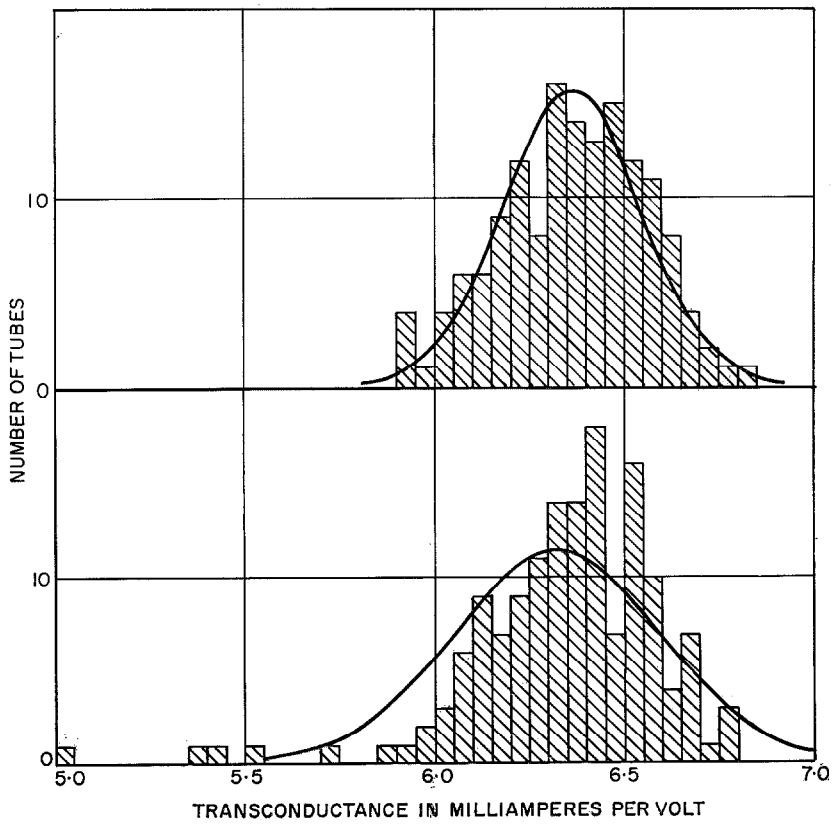


Figure 3—Histograms of transconductance of a group of *5A/182D* tubes. The lower curve is at the start of life test and the upper curve is after 1500 hours.

reached 4 years running time, shows a transconductance fall of 0·5 per cent per year over the past year. Additionally there is evidence that this already small rate of decline is diminishing with time.

6. Applications

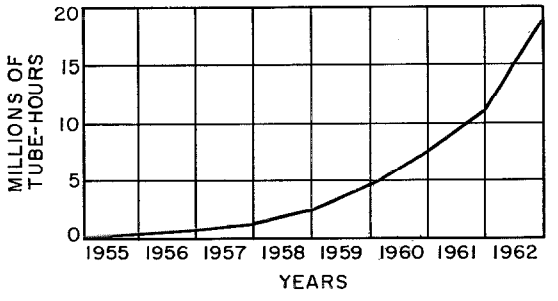
Examples of the tubes described in the foregoing are operating successfully in a large number of submarine-cable systems throughout the world. These include many installations around Europe. Half of the tubes installed in the repeaters of the United Kingdom-Canada cable laid in 1961 and the trans-Pacific cable now being laid between Vancouver, Canada, and Sydney, Australia, were manufactured in the plant surveyed above.

The chart in Figure 5 of tube-hours versus time continues to rise more steeply and it seems likely that the success of prior installations will encourage the planning and laying of wider-band systems with the newly developed tubes described above.

7. Acknowledgments

The writer is grateful for the help of all his colleagues in the submerged-repeater-tube section and especially to Mr. P. E. Douglas and Mr. N. B. Evershed. Thanks are also due to Mr. D. C. Rogers for some fruitful discussions during the course of preparation of this paper.

Figure 5—Cumulative tube-hours of tubes in submarine-repeater service from 1955 to 1962.



8. References

1. J. O. McNally, G. H. Metson, E. A. Veazie, and M. F. Holmes, "Electron Tubes for Transatlantic Cable System," *Proceedings of the Institution of Electrical Engineers*, volume 104, Part B, Supplement 4, pages 60–68; 1957.

2. G. H. Metson, "Study of Long-Term Emission Behaviour of an Oxide-Cathode Valve," *Proceedings of the Institution of Electrical Engineers*, volume 102, Part B, pages 657–675; September 1955.

3. M. F. Holmes and F. H. Reynolds, "Post Office Valves for Deep-Water Submarine Telephone Repeaters," *Proceedings of the Institution of Electrical Engineers*, volume 107, Part B, pages 165–171; March 1960.

4. E. J. Walsh, "Fine-Wire Type Vacuum Tube Grid," *Bell Laboratories Record*, volume 28, pages 165–167; April 1950.

5. G. T. Ford and E. J. Walsh, "Development of Electron Tubes for New Coaxial Transmission System," *Bell System Technical Journal*, volume 30, pages 1103–1128; October 1951.

9. Appendix

Consider a system comprising n repeaters in series. Each repeater contains two 3-tube amplifiers connected in parallel.

Let the probability of failure of any valve be p. Thus the probability of no failure $= (1 - p)$. The probability of no failure in any of 3 tubes $= (1 - p)^3$.

Thus the probability of failure

$$= 1 - (1 - p)^3$$
$$= p(3 - 3p + p^2)$$
$$= S.$$

With the parallel arrangement of amplifiers, the probability of failure of both channels

(which will fail the repeater) = S^2 and of no failure $1 - S^2$.

With n repeaters in series, the probability of the system not failing

$$= (1 - S^2)^n.$$

Thence the probability of breakdown of the system

$$= 1 - (1 - S^2)^n$$
$$= 1 - (1 - p^2(3 - 3p + p^2)^2)^n$$
$$\approx 9np^2.$$

Frederick Haegele was born in Hove, Sussex, England, on 1 May 1914. He obtained from the University College, Southampton, a B.Sc. degree in physics in 1934 and an M.Sc. degree in spectroscopy in 1937.

From 1937 to 1942, he was with Hanovia Ltd.; 1942 through 1945, with Mullard Radio Valve Co.; and during 1945 to 1950, with E. K. Cole Ltd.

He joined Standard Telephones and Cables in 1950 as production engineer on valves. Since 1955 he has been in charge of the development and production of submarine repeater valves

Mr. Haegele is an Associate of the Institute of Physics.

Ceramic-Insulated Vacuum Tubes for Very-High-Frequency Industrial Heating

J. J. BEHENNA
Standard Telephones and Cables Limited; London, England

The past decade has witnessed a remarkable expansion of the use of high-power tubes for industrial heating. Broadly speaking, this field can be divided into two distinct branches, induction heating and dielectric heating.

In induction heating the metal part to be heated is placed within or near a coil through which an alternating current is passed. A current induced in the metal produces heat as a function of the resistance offered to its flow.

In dielectric heating the material to be heated is placed between two metal plates with which it need have no physical contact. The two plates are connected to a radio-frequency voltage, and the material is heated by the current flowing through it as a function of its dielectric loss, which increases with frequency. Thus as frequency is increased, not only does the rate of heating increase, but materials having low dielectric losses can also be heated within reasonable times. The art of dielectric heating is still in the exciting stage where each new increase in power at a higher frequency heralds additional applications. Some present uses of dielectric heating are listed below.

(A) Pre-heating plastics and other materials to reduce production time and wear on tools.

(B) Welding of plastic materials, such as raincoats, inflatable toys, and upholstery.

(C) Packaging, as for shampoo packs.

(D) Wood-gluing.

(E) Unfreezing and melting such foods as chocolate, strawberries, and butter.

(F) Sterilizing of hospital packs.

(G) Disinfestation of pre-packed cake mixture, rice, flour, et cetera.

(H) Drying rolls and breakfast cereals.

(I) Drying textiles, veneers, paper products, bath cubes, foamed plastic and rubber, asbestos ceiling boards, and tobacco.

New uses for one of the tubes described in this article are as follows.

(A) Biscuit baking.

(B) Thawing bales of rubber frozen during transport through a very-cold climate.

(C) Thawing ice cream where a fault has occurred in processing.

(D) Thawing frozen fish.

(E) Sintering carborundum discs.

1. Tube Requirements

A new range of ceramic-insulated tubes have been designed for the highest frequencies now in use and to have adequate safety margins for operation at even higher frequencies for future applications. Other features are discussed below.

General mechanical robustness is essential because in the industrial heating field equipment will probably be used by unskilled operators and may be located in a machine shop, where vibration would be excessive for the normal radio types of tubes.

High safety factors against flashover across external terminals must be provided because a tube may be operated under very dirty and grimy conditions as in a wood shop or foundry and also in equipments where protection is not provided against voltage transients. The filament must be able to operate without mains regulation, as the filament voltage is not regulated in most industrial heating applications.

High efficiency is needed for economic reasons. In a 20-kilowatt-output radio-frequency generator, the cost of electric power represents about 75 per cent of the continuous running cost when depreciation and cost of tube replacements are taken into account.

Life must be not less than 5000 hours. In many cases tubes must be replaced by the manufacturer of the industrial heating equipment; a shut-down of a production line for tube replacement can be expensive.

The radio-frequency load may vary from full load to no load. This is reflected in the grid drive, where the grid dissipation under no-load conditions can be several times the dissipation under full load. The tube could fail if it is not designed to have an adequate grid safety factor.

The anode must be able to withstand momentary overloads.

2. New Range of Triodes

The range of tubes developed to meet these exacting requirements is given in Table 1.

To meet the high-frequency requirements these tubes have been designed with ceramic envelopes. Glass tubes have two main disadvantages, the dielectric loss in the glass at high frequencies and the overheating of the seals in handling the circulating radio-frequency currents, which can exceed 100 amperes. At these high frequencies the currents flow only through a very-thin skin of the conducting part, penetration generally being less than 0·001 inch (0·025 millimetre) at 100 megacycles per second. It is given by the expression

$$\mathrm{E} = k_1(R/\mu f)^{\frac{1}{2}} \qquad (1)$$

where E = skin depth

R = resistivity of the conducting material

μ = permeability of the conducting material

f = frequency in megacycles per second

k_1 = a constant.

The resistance of the conducting material is given by

$$R_{rf} = k_2(R/\mathrm{E}) = k_3(R\mu f)^{\frac{1}{2}}$$

where k_2 and k_3 are constants.

The radio-frequency circulating current is proportional to frequency, whence the radio-frequency losses P_{rf} are

$$P_{rf} = k_4 f^{\frac{3}{2}} R^{\frac{1}{2}} \mu^{\frac{1}{2}}$$

where k_4 is a constant.

In glass-to-metal seals, the metal is normally of a magnetic material, such as Kovar. As Kovar has high resistivity and permeability, the radio-frequency losses at the seals are therefore high, and at high frequencies cracking and/or glass suck-in near the seals can result.

With ceramic-to-metal seals this problem is minimized because the radio-frequency circulating currents at the seals flow through the metallizing and plating on the ceramic. The resistivity is low, and the permeability is unity.

Ceramics also have the additional advantages over glass of lower dielectric loss and higher safe operating temperature. The safe operating temperature of the seal may be between 220 and 250 degrees centigrade as against 180 degrees for Kovar-glass seals. Higher bake-out temperatures during evacuation increase both reliability and life.

Ceramic tubes will stand higher thermal and mechanical shocks than those with glass envelopes. They can be manufactured to more-exacting dimensional tolerances. Also, less chemical cleaning is required, since all brazes and final sealing-in can be carried out in a protective atmosphere.

TABLE 1
TUBE PERFORMANCE

Tube Type	Output Power in Kilowatts (Measured in Load)	Efficiency in Per Cent	Direct Anode Voltage	Maximum Frequency of Operation in Megacycles per Second
3JC/187E	5	80	6000	>220
3JC/203E	12	80	6000	220
3RC/223E	24	80	6000	>100
3ZC/262E	40	80	6000	100

TABLE 2

CHARACTERISTICS OF THE TUBES

	Length in Centimetres (Inches)	Amplification Factor μ	Mutual Conductance in Milliamperes per Volt
3JC/187E	21·8 (8·6)	12	22
3JC/203E	23·4 (9·2)	12	32
3RC/223E	26·2 (10·3)	12	60
3ZC/262E	36·5 (14·4)	12	60

TABLE 3

3JC/203E, TYPICAL MEASURED OSCILLATOR PERFORMANCE

Output Power, In Load	12 Kilowatts
Efficiency	80 Per Cent
Anode Direct Voltage	6 Kilovolts
Anode Direct Current	2·5 Amperes
Grid Direct Current	450 Milliamperes
Grid Bias Resistance	1500 Ohms
Peak Positive Grid Voltage	280 Volts

Tubes of the new range have smaller dimensions and lower inductances than previous designs, which contributes to efficiency at higher frequencies. Their tapered co-axial structures provide for easy insertion into high-frequency cavities.

The ceramic may be replaced by glass for induction-heating applications without other change of piece parts. Figure 1 shows the *3JC/187E* in both its glass and ceramic forms. The other three tubes are depicted in Figures 2 and 3.

3. Design of Metal-to-Ceramic Seals

The temperature coefficients of expansion of the metal and ceramic members of the seals must match as closely as possible, so that the seals can withstand repeated thermal cycling without fractures or vacuum leaks. Exact matching from ambient to brazing temperature is not possible with known materials, so careful attention must be given to the design of the seals. The high-alumina ceramic used has a tensile strength of 27 000 pounds per square inch (1898 kilogrammes per square centimetre) compared with a compressive strength of 10 times this value. The ceramic-to-metal seals have therefore been designed so that the ceramic is in compression after the brazing operation through the selection of a metal having a higher total thermal expansion than that of the ceramic and by brazing the metal components to the outside of the ceramic rings. The metal is a nickel-iron-

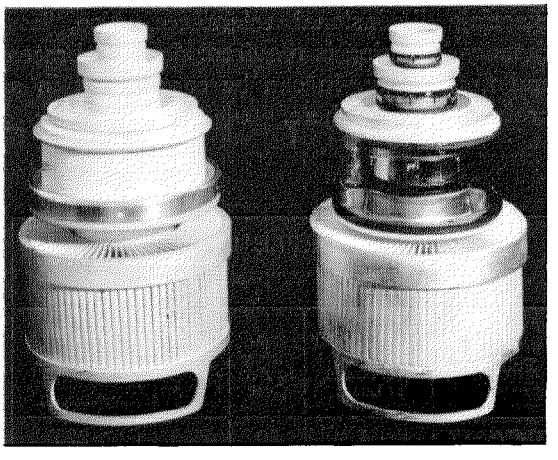


Figure 1—The ceramic and glass versions of the *3JC/187E* 5-kilowatt industrial triode.

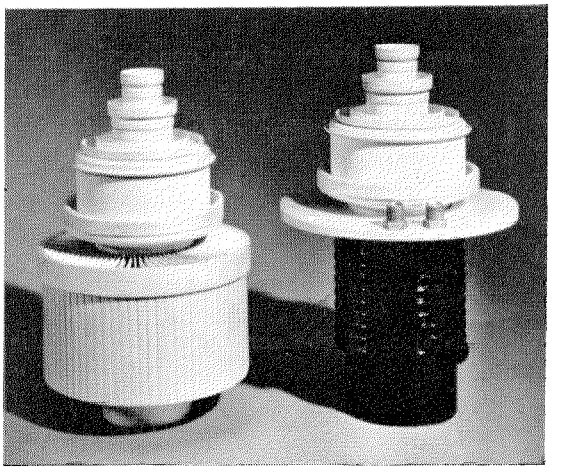


Figure 2—The *3JC/203E* 12-kilowatt and the *3RC/223E* 24-kilowatt industrial triodes.

cobalt alloy, and the ceramic is a 97·6-per-cent alumina.

4. Characteristics and Performance

4.1 Characteristics

To meet the important requirement of a high grid safety factor, the power dissipated in the grid must be kept low and the thermal emissivity of both the grid and anode surfaces should be as high as possible. To minimize grid power dissipation, the grid-to-cathode and grid-to-anode distances must be as short as mechanical stability and freedom from flash-over permit while a relatively coarse grid pitch, which in these tubes is about 1·4 times the grid-to-cathode distance, is used.

A more-detailed analysis of these design considerations is given in Section 12. The values of amplification factor and mutual conductance are given in Table 2.

The tubes are so similar in performance that only the *3JC/203E* will be examined in detail. Data for this tube are given in Table 3, and its characteristics in Figures 4 and 5.

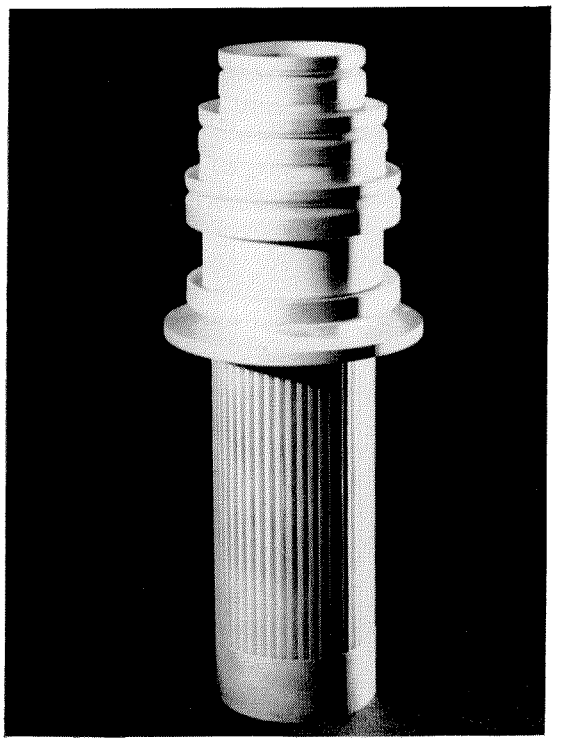


Figure 3—The *3ZC/262E* 40-kilowatt industrial triode.

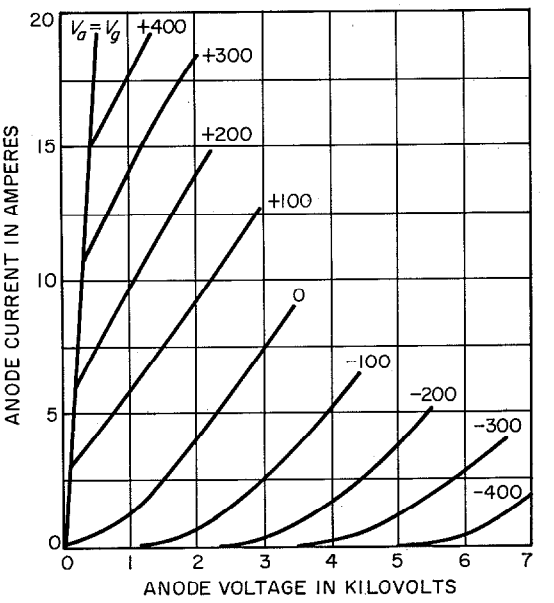


Figure 4—Anode current versus anode voltage for the indicated values of grid voltage for the *3JC/203E*.

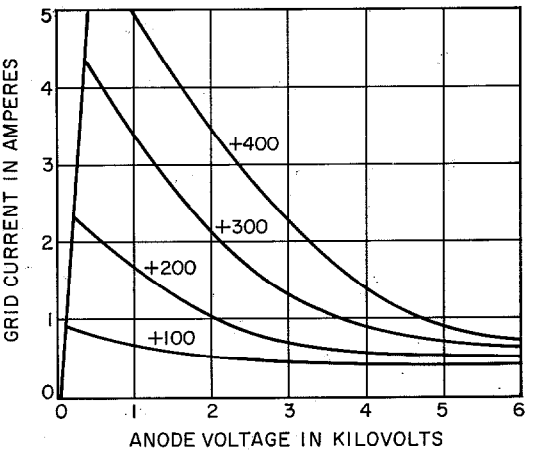


Figure 5—Grid current plotted against anode voltage for indicated grid voltages for the *3JC/203E*.

4.2 Efficiency

Using the symbols given in Section 10, the anode efficiency η is given by

$$\eta = \frac{(V - V_a)I_{1a}}{2VI_{da}}. \qquad (2)$$

For an anode-current flow angle of 140 degrees, which is representative of full-load operation, with a good compromise between adequate output power and efficiency [1], then $I_{1a}/I_{da} = 1{\cdot}77$, which substituted in (2), gives

$$\eta = 0{\cdot}885\,\frac{V - V_a}{V}$$

$$= 0{\cdot}885\left(1 - \frac{V_a}{V}\right). \qquad (3)$$

It can be seen from (3) that the maximum theoretical efficiency obtainable with a 140-degree flow angle is 88·5 per cent, when $V_a = 0$. In practice the anode "bottoming" voltage cannot be made as low as this because of the very-high grid dissipation that would result. However, the design principles that ensure a low value of grid dissipation with an adequate safety margin also allow a lower bottoming potential, and hence higher efficiency, than other available tubes of comparable power output. The *3JC/203E*, for example, can be operated with $V = 6$ kilovolts, $V_a = 570$ volts and $\phi = 140$ degrees, giving an anode efficiency of 80 per cent, with an adequate grid safety factor greater than 2:1. An even-greater efficiency could of course be obtained at the expense of power output and/or grid safety factor.

5. Filament Power and Emission

Among the earlier-listed requirements for industrial heating tubes is that the filament be able to operate without mains regulation and have a long emission life. These are, of course, conflicting requirements that must be weighed against each other in the design of the filaments.

5.1 Filament Temperature and Power

Thoriated-tungsten filaments in tubes for radio transmitting stations, where the filament voltage is normally controlled to better than ±1 per cent, are normally operated between 1950 and 2050 degrees kelvin. Emission increases with temperature, of course, for a given filament power, but a higher temperature shortens the emission life for a given wire diameter and carbide depth. To ensure adequate life within filament voltage variations of ±5 per cent these tubes were designed to operate at a low filament temperature at the nominal filament voltage. To ensure an adequate margin of emission under these conditions, it was calculated that the filament power must not be less than 38 watts per ampere of maximum useable emission. This adequate filament power also allows operation of the tubes with short-time variations of filament voltage as great as −10 per cent.

5.2 Carburization of Filaments

It is extremely difficult with high-frequency dielectric heating equipments, using voltage-step-up cavities to supply power to high-impedance loads, to prevent occasional circuit flashovers because of unfavourable atmospheric working conditions. In grounded-grid operation, external flashovers cause the anode voltage to appear momentarily on the grid, which can cause internal grid-to-filament flashovers, with consequent removal of the mono-molecular layer of thorium at the surface of the filament. Inability to replenish the thorium layer quickly can result in premature failure of the tube.

As is well known, a thoriated-tungsten filament is carburized to maintain at its surface a mono-molecular layer of thorium that, because of its low work function, provides the emission. The thorium results from the reduction of thoria in the following chemical equation:

$$2W_2C + ThO_2 \rightarrow 2CO + 4W + Th.$$

Several forms of tungsten carbide can exist, depending on the carburizing temperature and

pressure and on the nature of the carburizing gas. One of these forms of carbide reduces the thoria more readily than do the others. Hence, the carburizing conditions for the filaments of this range of tubes have been selected to provide this preferred carbide structure, so that the tubes can withstand the above-mentioned flashovers without loss of emission.

6. Mechanical Features

6.1 Envelope

A coaxial structure was chosen for its mechanical robustness and superior high-frequency performance over the more-conventional pin-type tubes. The coaxial structure leads to smaller dimensions and reduced inductances in the tube itself, thereby facilitating the use of these tubes in cavities for very-high-frequency operation. With exception of the *3ZC/262E*, all the tubes use the same stem parts, thus reducing tooling and jigging, multiplicity of parts, and costs. A cross-section of the *3JC/187E* showing details of the envelope structure is given in Figure 6.

The ceramic sub-assemblies, which are brazed in hydrogen at low pressure, are carefully jigged to ensure accurate concentricity and alignment of electrodes. No chemical cleaning is necessary after the brazing operations; thus, a major hazard of glass tubes, chemical residue leading to low emission, is eliminated.

The final seal-in of the tube is by argon-arc welding, thus avoiding any possibility of oxidization of the inside of the tube, which is a risk in glass-to-glass sealing.

If a dielectric-heating set employing a high-Q cavity is switched on without the load being connected, the peak radio-frequency voltage across the grid-to-anode ceramic can rise to several times the direct-current supply voltage. For this reason the metal parts at the grid and anode ceramic-to-metal seals have been carefully shaped to avoid field concentration and arcing across the anode-to-grid ceramic.

6.2 Filament

In all the tubes except the *3ZC/262E*, a mesh filament is used. It consists of a mesh of criss-crossed wires, wound on a cylindrical mandrel and welded together at the cross points. This type of filament resembles a uniformly emitting cylinder, thus enabling electrodes to be smaller than with more-conventional filaments, an important requirement for high-frequency operation. Another important feature is that the resistance to mechanical distortion is very great, which allows a smaller grid-cathode spacing. The filament is supported rigidly at one end, and at the other end is connected to a number of flexible tapes,

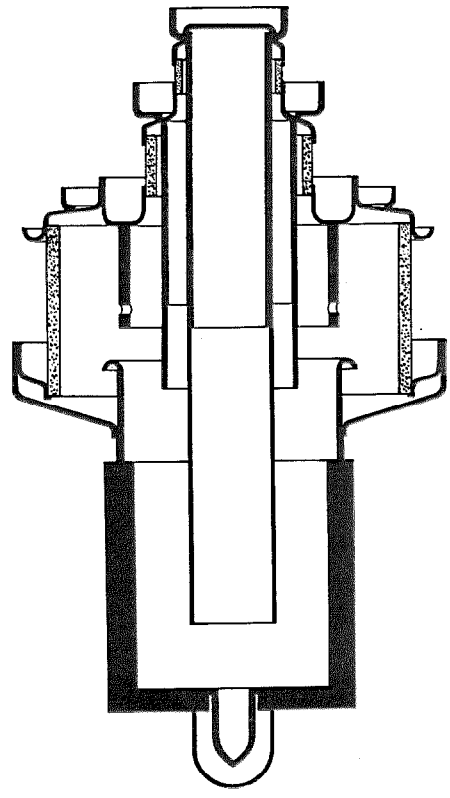


Figure 6—Cross section of the envelope structure of the *3JC/187E*.

as shown in Figure 7, which, in turn, are connected to the central tube that is used both for support and for conducting the filament current. The tapes allow expansion of the filament in the axial direction, thus minimizing mechanical strain.

The *3ZC/262E* has been designed with a long small-diameter anode to permit operation in a small boiler accommodated within a cavity. This shape does not readily lend itself to the mesh construction and so a more-conventional free-hanging filament is used with a slightly bigger grid-to-cathode spacing. Special design features allow for differential expansion between the individual filament strands and also prevent any barrelling of the filament in the radial direction. A special arrangement of the filament wires, in relation to the grid supports, allows an almost uniformly emitting cylinder to be obtained.

6.3 Grid

The grids of the mesh-filament tubes are made from expanded metal, specially processed to increase the thermal emissivity and to reduce grid electron emission to a very low value. This form of grid does not require grid supports; hence, it allows smaller electrode spacings.

The grid of the *3ZC/262E* is the conventional helically wound type, processed similarly to the expanded-metal grids. The grid and filament are mounted rigidly on a stem, but are fixed in a radial direction by a high-alumina bush that freely allows their differential axial expansion with respect to each other.

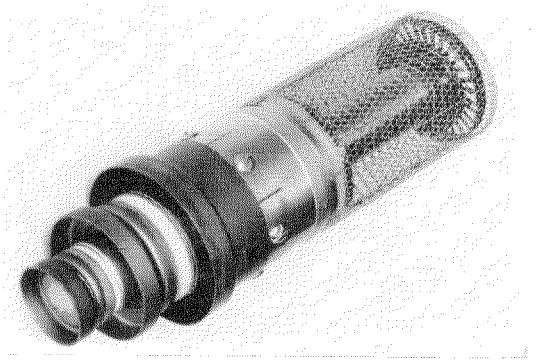


Figure 7—Grid-cathode assembly.

7. Cooling

Four different means of cooling are used.

(**A**) Air cooling (prefix J).

(**B**) Water cooling with normal water jacket (prefix Q).

(**C**) Water cooling with integral water jacket in the form of a coil hard soldered to the anode wall (prefix R).

(**D**) Vapour cooling (prefix Z).

All of the tubes, whatever the form of cooling, have thick-walled copper anodes to provide a large thermal overload capacity.

The air-cooled tubes have the fins brazed directly to the anode. This feature, coupled with the use of louvres on the fins to ensure turbulent flow, results in a high-efficiency radiator of minimum size. The *3JC/187E*, for example, requires only 50 cubic feet (1·42 cubic metres) per minute of cooling air at a water-gauge pressure of 0·4 inch (10·2 millimetres), for an anode dissipation of 3 kilowatts at an ambient temperature of 20 degrees centigrade.

Of the two forms of water cooling, the wrapped-coil version is finding increasing favour with manufacturers of industrial heating equipment. It is becoming the most favoured of all four cooling methods for medium-size equipments because of the very-small initial capital outlay for cooling apparatus, and also because of the small amount of cooling water required, an important feature where water is expensive. The required water flow is as low as that required for a vapour-cooled tube. The *3RC/223E*, for example, requires only 1 gallon (3·79 litres) per minute at a pressure of 7·5 pounds per square inch (0·53 kilogrammes per square centimetre) for 12 kilowatts of anode dissipa-

tion at a water inlet temperature of 20 degrees centigrade.

The advantages of vapour cooling are well known and will not be enumerated here.

The vapour-cooled versions of these tubes have vertical grooves machined into the anodes. This results in a large increase in the anode dissipation before heating occurs, and permits a much-greater anode dissipation than do the other water-cooled and the air-cooled versions. However the rated maximum anode dissipation of the vapour-cooled tubes has not been increased, and the advantage of vapour cooling is passed on to the customer in the form of an anode dissipation safety factor of over twice that of the normal safety factor for water- and air-cooled types.

8. Conclusions

The experience gained from the operation of a large number of glass tubes in a diverse range of industrial heating equipments, operating under many different adverse conditions, has been used in the design of a new range of improved ceramic-insulated tubes for this service.

Their high efficiency, generous safety factors, and high maximum frequency of operation provide a significant advance to the designer of very-high-frequency dielectric heating sets with output powers ranging from 5 to 40 kilowatts.

9. Acknowledgments

The author is greatly indebted to a large number of people who helped in the successful development of this range of tubes, especially to Messrs. H. M. Jenkins, G. H. G. Phipps, and A. R. Manners for their enthusiastic support and to Mr. D. C. Rogers for helpful criticism and encouragement.

10. Symbols

d = diameter of grid wire
I_a = peak instantaneous anode current per unit area
I_{da} = anode direct current per unit area
I_{dg} = grid direct current per unit area
I_g = Peak instantaneous grid current per unit area
I_k = Peak instantaneous total cathode current per unit area
I_{1a} = peak value of fundamental component of anode current per unit area
l_a = grid-to-anode distance measured to the centre of the wires
l_g = grid-to-cathode distance measured to the centre of the wires
p = pitch of grid wires
P_g = total grid dissipation per unit area of grid
R_b = grid-bias resistance
V = anode supply direct voltage
V_a = minimum instantaneous anode voltage relative to cathode
V_b = bias voltage
V_g = peak value of instantaneous grid voltage relative to cathode
V_s = peak value of grid voltage relative to bias voltage
η = efficiency
μ = triode amplification factor
ϕ = angle of grid-current flow in electrical degrees.

11. References

1. F. E. Terman, "Radio Engineering," McGraw-Hill Book Company, New York: 1932; pages 325 and following.

2. J. H. Fremlin, "Calculation of Triode Constants," *Electrical Communication*, volume 18, pages 33–49; July 1939.

3. J. H. Fremlin, R. N. Hall, and P. A. Shatford, "Triode Amplification Factors," *Electrical Communication*, volume 23, pages 426–435; December 1946.

4. W. J. Pohl and D. C. Rogers, "U.H.F. Triodes," *Wireless Engineer*, volume 32, pages 47–52; February 1955.

12. Appendix, Design Considerations

To establish design principles, expressions are required that relate tube electrode dimensions to the peak currents and voltages.

12.1 General Equations

An accurate expression [2] for the total space current is

$$I_k = \frac{2{\cdot}34 \times 10^{-6}\left(V_g + \frac{V_a}{\mu}\right)^{3/2}}{\left[l_g^{1/2} + \frac{(l_a + l_g)^{1/2}}{\mu}\right]^{3/2}},$$

amperes per square centimetre (4)

where voltages are expressed in volts and spacings in centimetres. Symbols are as given in Section 10.

For the grid current, the approximate relationship for a parallel-wire grid is

$$I_g = I_a\left(\frac{l_g + l_a}{l_g}\right)^{3/2}\frac{d}{p-d}\left(\frac{V_g}{V_a}\right)^{1/2}. \quad (5)$$

The grid flow angle ϕ is given by

$$\cos\frac{\phi}{2} = \frac{V_b}{V_s} = \frac{V_b}{V_b + V_g}. \quad (6)$$

A commonly used approximate expression [3] for μ is

$$\mu = \frac{2\pi l_a}{p\log_e\left(2\sin\frac{\pi d}{2p}\right)}. \quad (7)$$

Substituting (7) for μ in (4) and since in practice l_a is several times larger than l_g, then

$$I_k = \frac{2{\cdot}34 \times 10^{-6}\left(V_g + \frac{V_a K}{l_a}\right)^{3/2}}{[l_g^{1/2} + l_a^{1/2}K]^{3/2}},$$

amperes per square centimetre (8)

where $K = \dfrac{p\log_e\left(2\sin\frac{\pi d}{2p}\right)}{2\pi}$.

12.2 Grid Dissipation

An expression [4] for the grid dissipation is

$$P_g = \frac{\phi}{4\pi} V_g I_g. \quad (9)$$

One of the important tube requirements for industrial heating is a high grid safety factor, that is that P_g should be small.

From (9) it is evident that, to obtain this desirable feature, ϕ, V_g, and I_g should all be as small as possible.

The various electrode spacings and dimensions will now be examined to discover how the above requirements can best be met.

12.3 Grid-Anode Spacing

From (8) it is evident that for a given space current and current division l_a should be made small to keep V_g as low as possible. In other words the amplification factor should be made as small as possible consistent with an adequate safety factor as regards freedom from flashover. It might be argued that the resulting low amplification factor would necessitate a high drive power. This, however, is not true for the range of tubes described in this article because the drive power increases not only with a decrease in μ, but also with increases in both I_g and V_g, and the design features of these tubes ensure that V_g and I_g are small.

12.4 Grid-Cathode Spacing

Examination again of (8) shows that, for a given space current, l_g also should be made small to keep V_g as low as possible. In other words, the mutual conductance of the tube should be made as high as possible consistent with mechanical accuracy and stability of the filament and grid structures and supports.

It can be seen from (5) that for a given anode current, I_g also decreases as V_g is made smaller by achieving a low amplification factor and a high mutual conductance.

Further, from (6) it can be seen that a low value of V_g also results in a low angle of grid current flow ϕ.

Thus the above design considerations satisfy the requirement of low grid dissipation as shown in (9), that is, that V_g, I_g, and ϕ should all be small.

12.5 Grid Pitch

The larger the number of grid wires, the higher is the value of intercepted grid current for a given space current, resulting in a smaller anode current and, hence, power output. For a given amplification factor the grid safety factor is not affected because additional wires can dissipate additional power. The grid pitch should therefore be large, not only for the above reason, but also because this results in a smaller amplification factor, the benefits of which have been explained in 12.3 above. There is a limit, though, to increasing the grid pitch beyond a certain value because of the onset of "island" effects and "long tail" characteristics leading to low efficiencies. A grid pitch equal to 1·4 l_g is a good compromise [2].

13. Appendix, Grid Temperature and Emission

13.1 Runaway Conditions

The bias for an industrial heating oscillator is normally derived by the grid current I_{dg} flowing through a grid resistance R_g.

If in an oscillator an increase in the radio-frequency grid voltage $\int V_g$ is postulated, then resulting changes in the grid circuit will be:

(A) An increase in the normal grid current $\int I_{dg}$.

(B) An increase in the grid dissipation $\int P_g$.

(C) An increase in the grid emission current $\int I'_{dg}$.

The change in the bias voltage V_b is therefore given by

$$\int V_b = R_g[\int I_{dg} - \int I'_{dg}] \qquad (10)$$

and the tube will run away if

$$-R_g[\int I_{dg} - \int I'_{dg}] > \int V_g. \qquad (11)$$

If now a graph is plotted of $R_g[I_{dg} - I'_{dg}]$ against V_g with both ordinates to the same scale, then the tube will run away when the slope at that point is negative and greater than 45 degrees. It is interesting to note from this, that it is the slope of the grid emission curve that is important, not the absolute value of the grid emission. Since emission is related to temperature in such a way that the emission slope increases rapidly with temperature in accordance with the well-known emission law

$$I_s = A_0 T^2 \exp[-b_0/T] \qquad (12)$$

it is therefore very important that the grid temperature should be as low as possible.

13.2 Grid Temperature

Assuming that the power conducted along the grid and filament supports is small compared to the power radiated from the grid and filament to the anode, it can be shown that the grid temperature T_g is fairly accurately given by the expression

$$T_g^4 = \frac{P_g}{e_1 \sigma A_2} + \frac{P_f S}{\sigma A_2} + \frac{(P_g + P_f)(1 - e_2)S}{\sigma A_2 e_2} \qquad (13)$$

where A_2 = total surface area of grid

e_1 = total thermal emissivity of the grid

e_2 = total thermal emissivity of the inside surface of the anode

P_f = filament power

P_g = power received by the grid by electron bombardment
S = shadow area ratio of grid
σ = Stefan-Boltzmann constant.

Examination of (13) shows that the thermal emissivity of both the grid and anode surfaces should be as high as possible in order to keep the grid temperature low.

J. J. Behenna was born in Cornwall, England in 1927. In 1952, he received a Faraday House Diploma in Electrical Engineering from Faraday House Engineering College. He spent three years in the radio branch of the Royal Navy before going to college.

He joined Standard Telephones and Cables, New Southgate, in 1952, working on the development of high-power transmitters. Mr. Behenna, in 1957, transferred to the valve division at Paignton, and is now in charge of the high-power vacuum-tube development laboratory.

Travelling-Wave Tubes for 6-Gigacycle-Per-Second Radio Links

P. F. C. BURKE
Standard Telephones and Cables Limited; London, England

The recent development of the communication bands around 6 gigacycles per second has led to the design of two travelling-wave tubes for radio links. One of these, the *W5/1G*, operates at 3000 volts and 40 milliamperes for an operational output power of 5 to 8 watts. For 1800-channel systems of conventional design, the *W5/2G*, with an operating power output of 10 to 12 watts, is available. The required radio-frequency input power is about 1 milliwatt. Other important radio-frequency characteristics are noise factor, spurious noise, and both input and output matching.

A fairly high efficiency is necessary and is achieved by using a collector potential less than the helix potential. The reduced collector dissipation permits natural convection cooling. A periodic-permanent-magnet structure [1] requires no power for focusing the electron beam and is relatively small. The amplifiers comprise two parts: firstly, a mount with the waveguide feeds and magnets and, secondly, the travelling-wave tube. All parts subject to wear or failure are contained in the tube, which is easily replaced.

1. Design and Construction

The dimensions and other parameters of the tubes are given in Table 1 and tube and mount are shown in Figure 1. The helix is supported

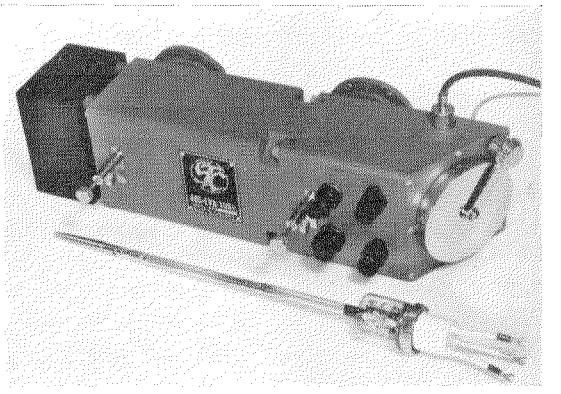


Figure 1—Travelling-wave tube *W5/1G* and mount.

TABLE 1
DIMENSIONS AND PARAMETERS OF TRAVELLING-WAVE TUBES

	W5/1G	*W5/2G*
Frequency Range in Gigacycles Per Second	5·85–7·05	5·85–6·45
Heater, Volts	6·3	6·3
Heater, Amperes	0·7	0·7
Grid-2 Volts, Approximate	2000	2300
Helix Volts, Approximate	2900	3400
Collector Volts	2000	2000
Collector Milliamperes	40	50
Helix Milliamperes, Approximately	0·5	0·5
Helix Inner Diameter, Inches (Millimetres)	0·088 (2·24)	0·088 (2·24)
Helix Wire Diameter, Inches (Millimetres)	0·012 (0·3)	0·012 (0·3)
Helix Pitch, Inches (Millimetres)	0·034 (0·86)	0·0345 (0·88)
Helix Length, Inches (Millimetres)	6·4 (163)	6·4 (163)
c/v	9·4	8·7
γa	1·62	1·5
F	0·39	0·39
Dielectric Loading Factor,* Per Cent	80	80
C	0·051	0·054
QC	0·2	0·2
Theoretical Synchronous Saturated Gain, Decibels	42	40
Helix Wire Loss, Decibels	6	6
Central Loss, Decibels	70	70
Cathode Diameter, Inches (Millimetres)	0·2 (5·1)	0·2 (5·1)

* Calculations based on Pierce [2] and Tien [3].

by 3 insulating rods with a carbon deposit to give a central attenuating region. A fusible link enables the helix to be outgassed by passing current through it on the pump. The helix pitch must be accurately maintained; the turn-to-turn spacing is checked with an image-splitting eye-piece [4] ensuring that errors are less than 1 per cent of the pitch. The electron gun is assembled on ceramic rods and is supported between the helix assembly and a metal ring at the base of the tube. The first grid of the electron gun is connected internally to the cathode. The beam current is controlled by the second grid. The third grid, also connected internally to the helix, has an external connection in the base ring. Ceramic lead-through insulators provide for connections to the electrodes and permit pumping through a relatively large copper tubulation in the metal base. Lead-throughs and tubulation are protected by potting in Silastomer after final processing of the tube. Individual components are heated in vacuum whenever possible before assembly to minimize the evolution of gas during life.

2. Mount

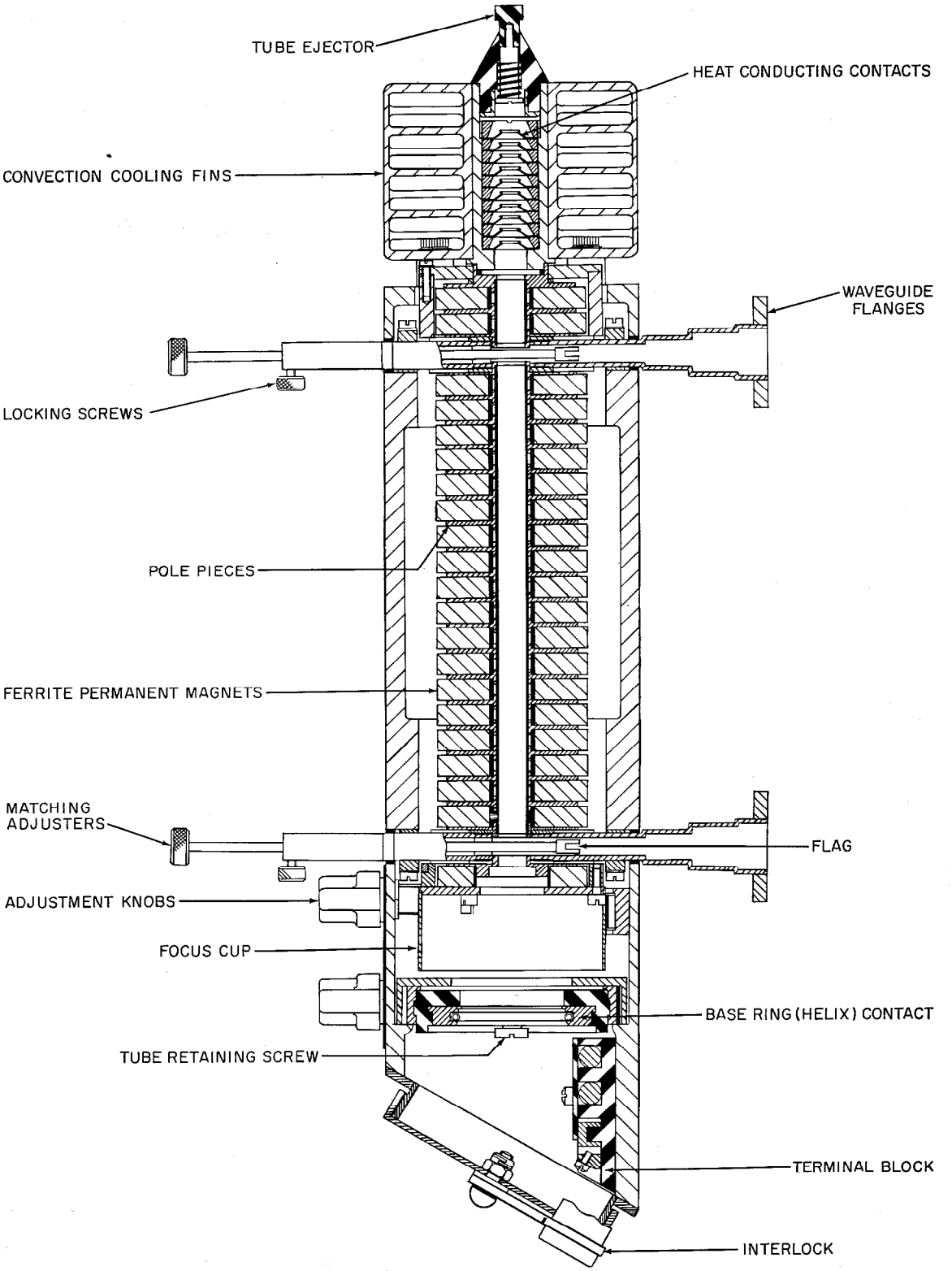
The mount of the *W5/1G* is shown diagrammatically in Figure 2.

Adjacent permanent magnets have opposite polarities, greatly reducing the leakage field and the total flux compared with that which would be obtained from a magnet providing a uniform field. The volume and weight are reduced accordingly. Short magnets are necessary to avoid defocusing due to resonant interaction with the electron beam and hence the pole-pieces need to have a small internal diameter to give the required field. This resonance will still occur at low voltages and leads to a rapid deterioration in the focusing at voltages appreciably below the normal operating voltage. Bar magnets, not shown in the diagram, provide the field across the waveguides and are mounted with the central magnet stack in an aluminium casting of rectangular external cross-section. The cooler, attached so as to conduct little heat back to the mount, has cooling fins of aluminium with a total surface area of approximately 120 square inches (774 square centimetres). Good thermal contact between the rigidly mounted cooler and the collector of the tube is provided by approximately 80 spring fingers pressing flexibly against the collector. Silver-bearing copper, 0·010 inch (0·25 millimetre) thick, is used since it does not set under stress at temperatures below 200 degrees centigrade. Since each finger gives two or three points of contact the total heat drop between the collector and the interior of the cooler is less than that obtained with a normal mechanical fit between the collector and a solid cooler. For 80 watts of dissipation, this temperature drop is about 10 degrees centigrade and the cooler itself runs about 70 to 80 degrees centigrade above ambient temperature. The fins of the cooler are staggered to increase the turbulence of the ascending air and side pieces are incorporated to give a chimney effect. The tube can be freed from the grip of the spring contacts by an ejector that applies pressure to its end.

The base ring of the tube is held at the other end of the mount in a ring whose position can be altered by two knobs to correct for any slight errors in alignment of the base. Two other knobs move the first pole-piece to give variable transverse magnetic fields in the region of the electron gun. Both adjustments are limited to prevent any misalignment that would damage the tube. Electrical adjustment of the current in special deflector coils is an alternative to movement of the pole-piece.

The waveguide in the mount is 1·375 by 0·25 inch (3·5 by 0·64 centimetre) with a binomial matching section to *WG137* dimensions. Some fine trimming is necessary to obtain the best match at both input and output either with an adjustable flag, as shown in Figure 2, or with

Figure 2—Structural details of the mount. ⟶


TUBE EJECTOR
HEAT CONDUCTING CONTACTS
CONVECTION COOLING FINS
WAVEGUIDE FLANGES
LOCKING SCREWS
POLE PIECES
FERRITE PERMANENT MAGNETS
MATCHING ADJUSTERS
FLAG
ADJUSTMENT KNOBS
FOCUS CUP
BASE RING (HELIX) CONTACT
TUBE RETAINING SCREW
TERMINAL BLOCK
INTERLOCK

four matching screws between the waveguide flange and the mount.

The flexible leads from the tube are connected to screwed terminals just inside the mount. Resistors are provided in the helix and grid-2 leads from these terminals to limit surges in the unlikely event of momentary voltage breakdown in the tube. The leads have polytetrafluoroethylene insulation and are taken, via a ferrite that absorbs any radio-frequency energy, through a cable whose outer braid is connected to the mount and should be earthed.

A lid, which screws to the mount, gives good radio-frequency screening. On removing the lid, a plug swinging on an attached arm is disconnected from its socket in the mount and removes all high voltages.

The higher-power tube *W5/2G* uses a slightly modified mount incorporating mechanical adjustment of the tube position at the collector end.

3. Electron Beam Focusing

The magnetic field distribution in the periodic structure is shown in Figure 3. Slight increases in period at the waveguides avoid the use of extremely narrow guides. A gradual increase in magnetic field over the last four periods before the output waveguide compensates for defocusing effects at maximum power output. The values of the first two peaks of magnetic field were found experimentally to give best focusing. The magnetic field gradually falls in the region of the electron gun until it is zero at the cathode.

The ferrite permanent magnets used have an appreciable temperature coefficient but the helix current does not rise by more than 50 per cent over the temperature range from −10 to +60 degrees centigrade. The materials and the working point of the ferrite are such that exposure to temperatures of −60 degrees centigrade gives no irreversible change in the magnetic field.

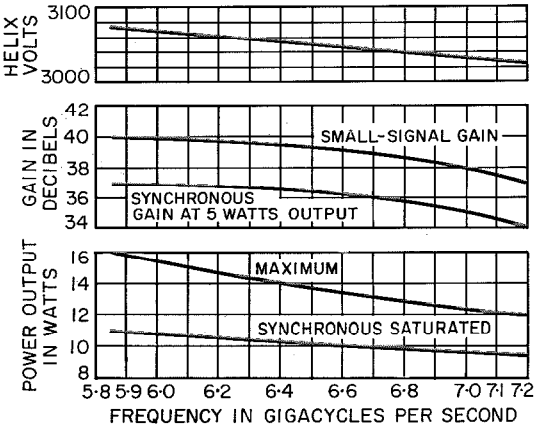


Figure 4—Gain, power output, and helix voltage for synchronous operation as a function of frequency for the *W5/1G*.

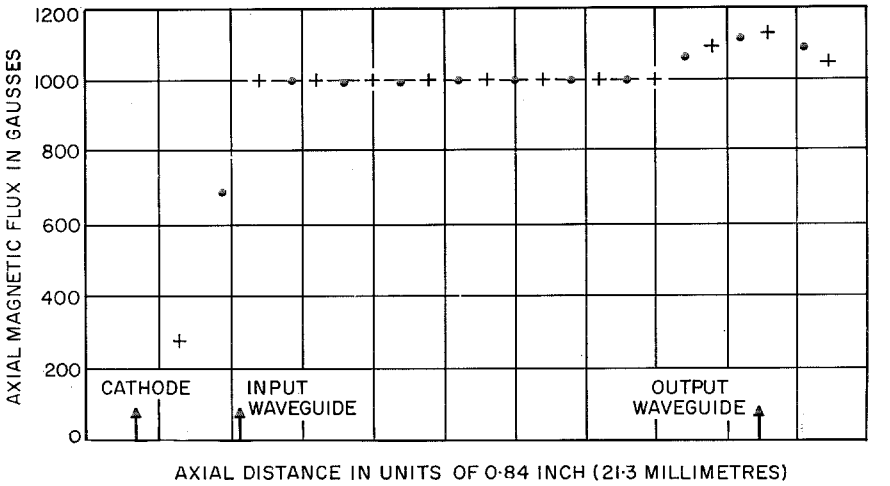


Figure 3—Magnetic field distribution in the periodic-permanent-magnet structure.

Periodic focusing is not effective at low voltages, which means that precautions are necessary on first switching on. If the beam current is held off by keeping the second grid at cathode potential until the desired helix voltage is set up, the current surge as V_{g2} rises to the normal operating level will be less than 10 milliamperes. If it is kept a constant fraction of the helix voltage, the current surge may be 15 milliamperes. No excessive voltage drop, due to power supply impedance, should occur with these surges.

Measurements of helix current at maximum power output on a considerable number of tubes and mounts gave a figure of less than 0·5 milliampere for half the results and 1 milliampere was exceeded in only 2 per cent of the cases. The normal operating maximum in field use is specified as 1·5 milliamperes initially and the supplies should be removed if 2·5 milliamperes is exceeded.

4. Radio-Frequency Performance

4.1 Gain and Power Output

Gain and power output are shown as functions of frequency for the *W5/1G* and *W5/2G* in Figures 4 and 5. Typical graphs of output power versus input power for three different helix voltages at a frequency of 6·4 gigacycles per second are shown in Figure 6 for the *W5/2G*. Higher helix voltages give greater saturated output but the tubes are designed

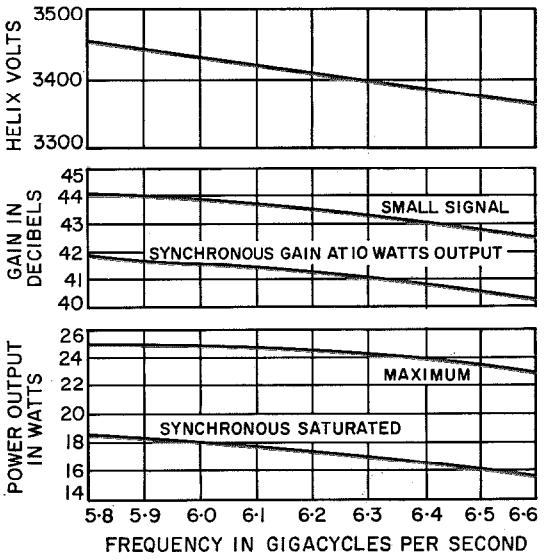


Figure 5—Gain, power output, and helix voltage for synchronous operation as a function of frequency for the *W5/2G*.

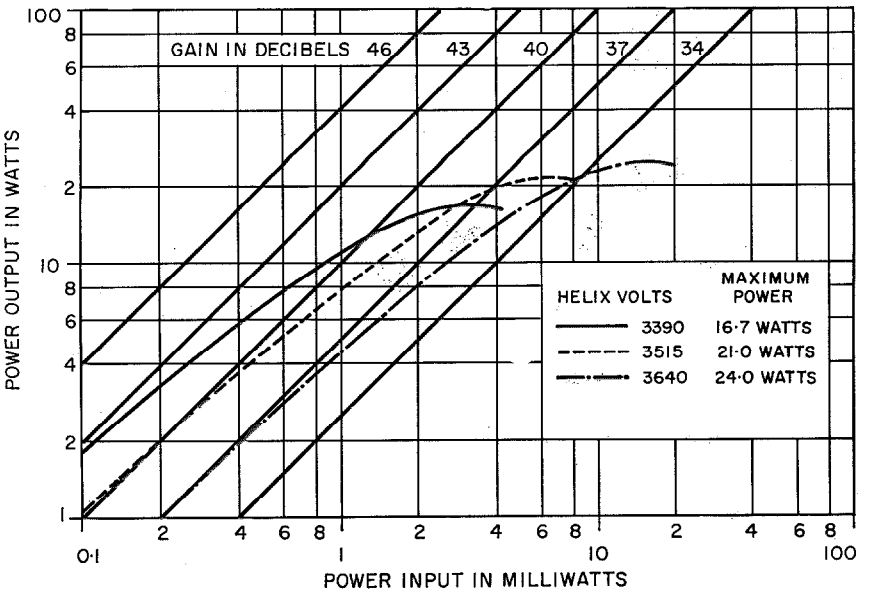


Figure 6—Output power plotted against input power for *W5/2G* at 6·5 gigacycles per second.

to maximize the power obtained at the helix voltage giving maximum low-level gain, so as to give less distortion from the conversion of amplitude modulation to phase modulation.

A slight fall in the saturated output power may be seen sometimes immediately after switching on the radio-frequency input. This lasts a minute or so and is due to parts of the tube increasing in temperature and in radio-frequency loss. For the *W5/1G* this fading gives a 12-per-cent power drop at a saturated output of 15 watts and 10-per-cent at 10 watts but at the same power level the *W5/2G* has a drop of only from 15·5 watts to 15 watts, perhaps due to the use of fused silica rods instead of steatite rods as in the *W5/1G*.

4.2 Signal–to–Noise Ratio

The noise factor of both tubes lies in the range of 26 to 28 decibels. Experiment has shown this to be independent of the exact magnetic field at the cathode over the range from 0 to 30 gausses, the discrepancy with earlier reports of high noise factor at zero cathode fields [5] being due perhaps to the way the field extends close to the cathode and controls the beam over almost all of its path.

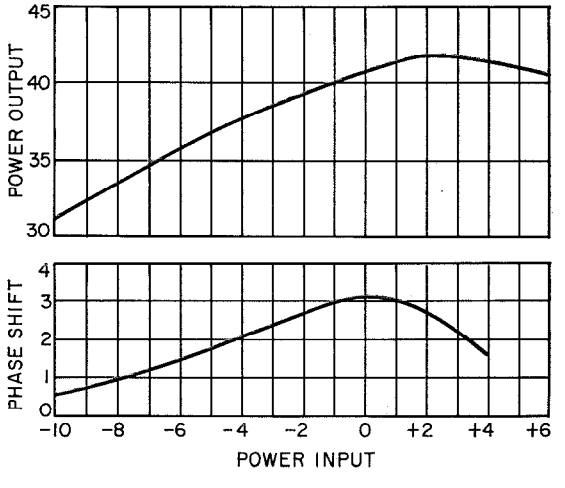


Figure 7—The effect of power input on power output, both in decibels referred to 1 milliwatt, and on phase shift in electrical degrees per decibel of input power are plotted above for the *W5/1G*.

The modulation noise due to ion-plasma oscillations has been studied and considerably reduced by attention to processing of the tubes. Sidebands between 20 kilocycles per second and 100 kilocycles per second wide at frequency distances around 3 megacycles per second and 10 megacycles per second from the carrier may initially have an amplitude of 10 to 20 decibels above the normal noise level but this falls rapidly during activation and testing to a level about 5 decibels above the shot noise. Improvement continues until no noise is detectable, although during the first 50 hours slight alterations in the operating conditions may make it perceptible again.

4.3 Leakage and Attenuation

Screening is good enough to ensure at least 65 decibels of attenuation of leakage between either of the waveguides and a radiating waveguide horn held in any position 2 inches (5 centimetres) or more from the mount.

Reverse attenuation, measured from tube output with the tube running at normal voltage and current, exceeds 70 decibels, effectively the same as the cold attenuation. Some reflected electrons from the collector decrease this value slightly with the helix about 200 volts below the normal operating value, corresponding to synchronism with a reflected electron beam of very low current.

4.4 Matching

For best results, matching adjustments should be made at the operating frequency when a new tube is inserted in a mount, giving a bandwidth of greater than 15 megacycles per second for a reflection coefficient of 5 per cent looking into the output of the tube, the limit being set by reflections at irregularities in the helix, which are amplified by the electron beam. At the operating level the gain is lower and the bandwidth improved. The bandwidth of the input match is considerably better, because of the lower gain before the attenuation region.

Alternatively, mounts can be pre-set for oper-

ation in any specified 600-megacycle-per-second range to give a voltage standing-wave ratio of less than 1·6 on all tubes without adjustment.

4.5 Phase Characteristics

The electrical length of the *W5/1G* is 30 wavelengths at 6 gigacycles per second and the *W5/2G* is 28 wavelengths long. Phase sensitivity is about −0·75 degree per volt change of helix potential and +0·125 degree per volt change of the second-grid potential at the normal operating power level. No measurable variation with collector voltage exists. The phase shift also varies with input power because of the slowing down of the electron beam as more power is taken from it. The differential of this quantity, expressed in degrees per decibel, is shown in Figure 7 for the *W5/1G* at synchronous voltage. At higher voltages, giving increased output, the maximum value can be 6 degrees per decibel.

4.6 Harmonic Output

Non-linearity near maximum power output produces many harmonics from travelling-wave tubes although this output may vary considerably since the impedance match at the harmonic frequencies will be generally bad and erratic. Second-harmonic power from the *W5/1G* at saturation output can be as much as 50 to 100 milliwatts.

5. Life

Cathode-current density is a most-important factor in determining life. Previous designs of medium-power travelling-wave tubes have given average lives of 7000 to 8000 hours at a cathode-current density of 400 milliamperes per square centimetre, and around 20 000 hours at 200 milliamperes per square centimetre. The *W5/1G* operates at this lower figure and the *W5/2G* at a value about 20 per cent higher. Hence a life of 15 000 to 20 000 hours is expected though, as this represents several years of running, it has not been verified yet.

Automatic control of the cathode current by the second-grid voltage is possible as with previously developed tubes. This control may not be necessary if maintenance visits are possible every few weeks since these will permit manual adjustments during the last part of the tube life.

The cathode temperature should be carefully controlled by regulating the heater voltage to within ±3 per cent of the recommended value of 6·3 volts, although short-term excursions as large as ±5 per cent are permissible. This establishes the optimum condition between any accelerated aging at high temperatures and irreversible poisoning at low temperatures.

6. Conclusions

Two travelling-wave tubes suitable for use in 6-gigacycle-per-second links have been developed. Although primarily developed for a superheterodyne type of microwave repeater, they can be used in other types as well [6]. Tubes giving similar performance in the adjacent higher frequency bands up to 8·5 gigacycles per second are in an advanced state of development.

7. Acknowledgments

This development could not have been successful without considerable and unstinted efforts from all of those concerned. Their numbers necessitate their anonymity, and the author's departure from this is solely an acknowledgment of responsibility for the shortcomings of this paper.

8. References

1. J. T. Mendel, C. F. Quate, and W. H. Yocom, "Electron Beam Focusing with Periodic Permanent Magnet Fields," *Proceedings of the IRE*, volume 42, pages 800–810; May 1954.

2. J. R. Pierce, "Travelling-Wave Tubes," D. Van Nostrand, New York, 1950.

3. P. K. Tien, "Traveling-Wave Tube Helix Impedance," *Proceedings of the IRE*, volume 41, pages 1617–1623; November 1953.

4. J. Dyson, "Precise Measurement by Image Splitting," *Journal of the Optical Society of America*, volume 50, page 754; 1960.

5. P. F. C. Burke and D. C. Rogers, "Magnetic Focusing of Long Cylindrical Electron Beams," *L'Onde Electrique*, volume 37, pages 174–189; February 1957.

6. C. I. McDowell, "An S. H. F. High-Density R. F. Heterodyne Repeater Microwave System." Presented at 1963 Winter General Meeting of the Institute of Electrical and Electronics Engineers.

P. F. C. Burke was born in London, England, on 16 October 1926. He obtained a B.A. degree in physics in Part 2 of the Natural Science Tripos from Trinity College, Cambridge, in 1948 and a M.A. degree in 1962.

In 1948, he joined Standard Telephones and Cables Limited and is engaged in the development and engineering of travelling-wave tubes.

Mr. Burke is a graduate member of the Institution of Electrical Engineers.

Effect on an Electron Beam of Variations in Periodic Permanent-Magnet Focusing Systems *

B. MINAKOVIC
Standard Telephones and Cables; London, England

The theory and design of a periodic permanent-magnet focusing system are based on the assumption that both the period and peak magnetic field are constant. This ideal, however, is never attained in practice. Non-uniformities in the period or magnetic field may be due to various tolerances or to the mechanical design of the circuit. On the whole, undesirable variations tend to increase the beam ripple which, in the case of the travelling-wave tube, may lead to loss of gain, increased frequency sensitivity, and high intercepted current.

In this paper, the effects of various practical non-uniformities in the period and magnetic field are examined by calculating the appropriate beam profiles.

The effect of small random misalignments of magnetic cells is also considered, and a simple expression is given for the total beam displacement from the intended circuit axis.

1. Beam Profiles for Non-Uniform Period or Magnetic Field

1.1 Outline of Analysis

The beam profiles were obtained by solving numerically Harker's equation [1] with a more-general expression for the axial magnetic field.

$$\frac{d^2\sigma}{dx^2} + 2\alpha B^2\sigma - \frac{2\alpha K}{\sigma^3} - \frac{\beta}{\sigma} = 0 \qquad (1)$$

where $\alpha = \dfrac{\eta L^2 \hat{B}_o^2}{64\pi^2 V}$

$$\beta = \frac{\rho_o L^2}{16\pi^2 \epsilon_o V}$$

$$K = \left(\frac{r_c}{r_o}\right)^4 \left(\frac{B_c}{\hat{B}_o}\right)^2$$

$$\sigma = r/r_o.$$

The suffix c indicates the values at the cathode plane, and o is the values at the plane of entry into the focusing circuit. V is the beam voltage, ρ is space-charge density, and r is the beam radius.

Equation (1) is derived by starting with three equations of motion in cylindrical co-ordinates. The radial force due to space charge is obtained from Gauss's theorem and this is reasonably correct provided that the envelope slope is less than, say, 0·15. Since the beam moves in a unipotential tunnel, the axial acceleration is negligible and the z equation can be ignored. With cylindrical symmetry in the magnetic field, the θ equation can be integrated, leading finally to (1).

Instead of commonly used

$$B = \cos x$$

the axial magnetic field is written here as

$$B = \bar{B}_n \cos\{l_n(x - X_n) + (n - 0{\cdot}5)\pi\} \qquad (2)$$

where $X_{n+1} = \dfrac{\pi}{l_n} + X_n$

$$X_o = -(\pi/2l_o)$$

$$l_n = L/L_n$$

$$\bar{B}_n = B_n/\hat{B}_o$$

and

L = optimized period of a uniform stack

L_n = period of the nth magnetic cell in a non-uniform stack

n = magnetic-cell reference number

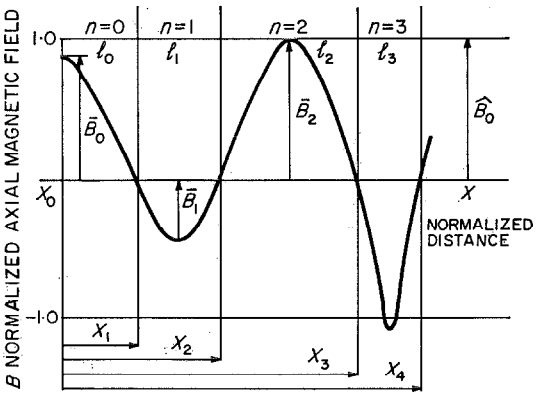
$\hat{B}_o$ = optimized peak magnetic field of a uniform stack

B_n = peak magnetic field of the nth magnetic cell in a non-uniform stack.

Equation (2) represents a cosine curve, which may have an arbitrary amplitude and/or period in any interval between two zero values, as shown in Figure 1.

* Presented in part at the International Congress on Microwave Tubes at Scheveningen, Holland, in 1962.

Figure 1—Cosine function of variable period and amplitude distribution of magnets.



The usual initial conditions were used, that is at $x = 0$, $\sigma = 1$ and $d\sigma/dx = 0$. Unless otherwise stated, α and β are the optimized values.

1.2 Random Variations in Period

Small random variations in the period of a focusing circuit are caused by the tolerances in magnets, pole-pieces, and circuit assembly. They will also occur when the magnetic-field distribution does not coincide with the cell geometry.

Variations of ± 1 per cent about the optimized period L were considered. Beam profiles were calculated for three different assemblies, each consisting of 15 magnets. The periods expressed in terms of l_n are given in Table 1, where l_n is the ratio of optimized period to actual period; and the maximum increase in beam radius is given in Table 2, where $\Delta\sigma$ is the maximum percentage increase in the normalized beam radius above unity.

It is interesting to note that for small values of α and β the increase in beam radius is of no practical importance.

1.3 Random Variations in Peak Magnetic Field

The field within a magnetic cell of an assembled periodic-permanent-magnet focusing stack is the result of the superposition of the fields in a cell alone, and of the stray fields of the neighbouring cells. Due to inhomogeneities in magnets and pole-pieces, which cause variations in stray fields and in the reluctance of the various paths, the peak magnetic field in an assembled stack is seldom absolutely uniform.

Again, as in the last section, three stacks were considered. As before, each stack had 15 magnets, but this time the random variation was obtained from actual measurement on experimental circuits. Table 3 gives the

TABLE 1
Variations in the Period

l_n	Assembly		
	1	2	3
l_0	1	1	0·9901
l_1	1·01	1·01	1·01
l_2	1	1·01	1·01
l_3	0·9901	1	0·9901
l_4	0·9901	1·01	1·01
l_5	1·01	0·9901	1
l_6	0·9901	1·01	1
l_7	1·01	0·9901	1
l_8	0·9901	0·9901	0·9901
l_9	1·01	1	0·9901
l_{10}	1·01	1·01	1
l_{11}	0·9901	0·9901	1
l_{12}	0·9901	1·01	1
l_{13}	0·9901	1·01	1
l_{14}	1	1·01	1

TABLE 2
Per Cent Increase in Normalized Beam Radius for Period Variation in Table 1

K	α	β	$\Delta\sigma$ in Per Cent		
			Assembly 1	Assembly 2	Assembly 3
0	0·1	0·0959	0·23	0·15	0·12
	0·3	0·2630	2·1	2·2	1·4
	0·5	0·3893	12·6	5·7	5·0
0·1	0·1	0·0749	0·33	0·35	0·22
	0·3	0·1857	2·3	2·9	1·2
	0·5	0·2129	5·1	4·6	5·1

TABLE 3

VARIATIONS IN THE PEAK MAGNETIC FIELD

Normalized Peak Magnetic Field $\bar{B}_n$	Assembly		
	1	2	3
$\bar{B}_0$	0·966	1·050	1·037
$\bar{B}_1$	0·915	0·973	0·954
$\bar{B}_2$	1·034	0·990	1·037
$\bar{B}_3$	0·983	0·994	1·037
$\bar{B}_4$	1·051	1·006	0·986
$\bar{B}_5$	1	0·99	1·012
$\bar{B}_6$	1·027	1	0·981
$\bar{B}_7$	1·017	1·017	0·996
$\bar{B}_8$	1·017	1·006	0·964
$\bar{B}_9$	1·003	1·006	1·020
$\bar{B}_{10}$	1·027	1·006	0·986
$\bar{B}_{11}$	1·017	1·006	1·008
$\bar{B}_{12}$	1·010	1·013	0·974
$\bar{B}_{13}$	1·003	1·006	1·06
$\bar{B}_{14}$	0·949	0·957	0·981

TABLE 4

PER CENT INCREASE IN NORMALIZED BEAM RADIUS FOR FIELD VARIATION IN TABLE 3

K	α	β	$\Delta\sigma$ in Per Cent		
			Assembly 1	Assembly 2	Assembly 3
0	0·1	0·0959	12·3	5·4	9·3
	0·3	0·2630	39·3	13·6	30·8
	0·5	0·3893	27·8	37·3	38·6
0·1	0·1	0·0749	13·1	6·2	8·7
	0·3	0·1857	85·3	13·2	42·8
	0·5	0·2129	27·0	11·5	44·2

normalized random values of the peak magnetic field, and Table 4 gives the corresponding increases in the beam radius.

For the comparison with a 1-per-cent variation in the period, beam profiles were also calculated for a 1-per-cent variation in the peak magnetic field. The values of $\bar{B}_n$ are now identical with l_n in Table 1, Assembly 1.

The relatively large percentage increases in the beam radius shown in Tables 4 and 5 are somewhat unexpected, as it is sometimes

TABLE 5

PER CENT INCREASE IN NORMALIZED BEAM RADIUS FOR A 1-PER-CENT PEAK MAGNETIC-FIELD VARIATION IN TABLE 1 ($l_n - \bar{B}_n$, ASSEMBLY 1)

K	α	β	$\Delta\sigma$ in Per Cent
0	0·1	0·0959	7·7
	0·3	0·2630	20·1
	0·5	0·3893	28·9
0·1	0·1	0·0749	5·4
	0·3	0·1857	12·3
	0·5	0·2129	9·2

believed that relatively large variations in peak magnetic field are quite tolerable.

It was thought that this discrepancy was caused by the operating α being always larger than its theoretical optimized value. For this reason some of the above computations were repeated with $\alpha = 2{\cdot}5\ \alpha_{opt}$; the lower curve in Figure 2 appears to confirm this. The same reasoning should apply to the effect of variations in the period.

1.4 SINGLE INCREASE IN PERIOD

The radio-frequency coupling to a travelling-wave tube can be obtained via a helical coupler, a radial or co-axial cavity, or by a waveguide transducer. Unlike a helical coupler, which may be small enough to be inserted inside the pole-pieces, a waveguide or cavity must be inserted between the magnets.

It is possible to retain the correct field across the waveguide, either by using additional bar magnets or utilizing the stray fields from the adjacent ring magnets.

The height of a standard waveguide section is usually much greater than half of the necessary period, and consequently, to avoid an increase in the period, the waveguide is stepped down. Very often drastic reduction in the waveguide height makes radio-frequency matching unsatisfactory; and then an increase in the period is unavoidable.

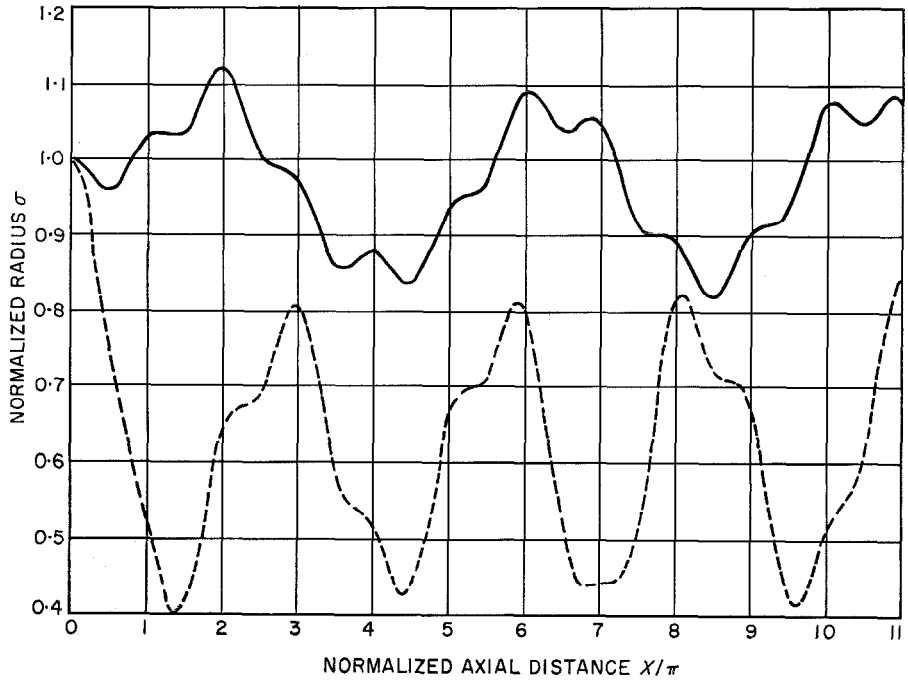


Figure 2—Beam profiles for random variations in the peak magnetic field of assembly 1 of Table 3 with $K = 0$ and $\beta = 0{\cdot}0959$. The solid line is for optimum $\alpha = 0{\cdot}1$ and the broken line is for $\alpha = 0{\cdot}25$.

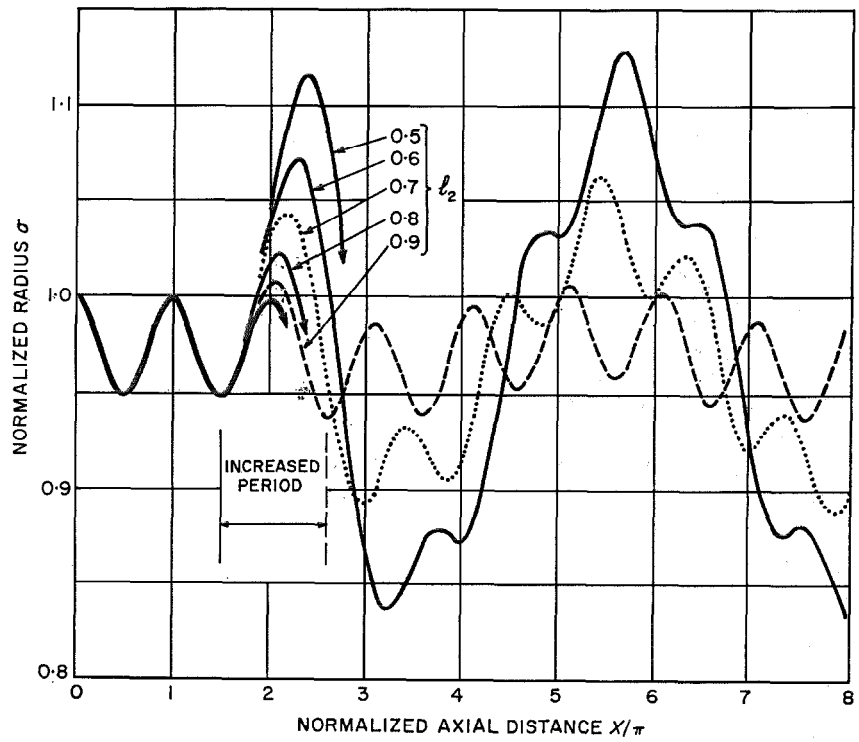


Figure 3—Beam profiles for a single increase in period for $K = 0$, $\alpha = 0{\cdot}1$, and $\beta = 0{\cdot}0959$.

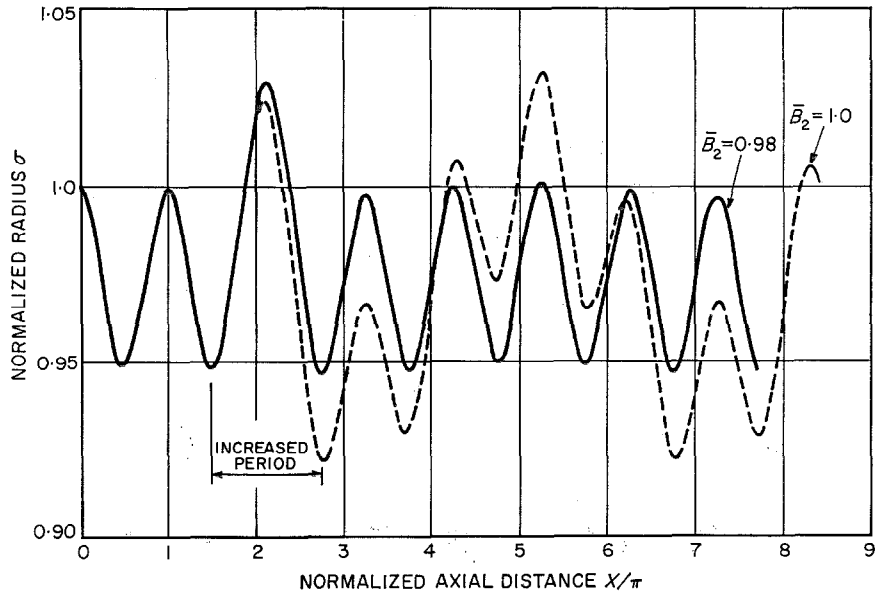


Figure 4—Beam profile before and after compensation for a single increase in period for $l_2 = 0{\cdot}8$, $K = 0{\cdot}1$, $\alpha = 0{\cdot}1$, and $\beta = 0{\cdot}0749$.

The effect of such a single increase in the period is shown in Figure 3. The usual sinusoidal beam profile is considerably distorted as l_2 becomes smaller, that is, the gap is increased. If, however, an increase in the period is accompanied by an appropriate reduction in magnetic field, the beam profile outside the waveguide remains undisturbed. Thus, for a particular set of parameters as in Figure 4, the correct field across the waveguide is $0{\cdot}98\ \hat{B}_o$.

According to Figure 3, doubling the period, that is, $l_2 = 0{\cdot}5$, causes only about 15-per-cent increase in the maximum beam radius in the middle of the gap. Although this would be acceptable in practice, experience shows that if the period increase exceeds about 20 per cent, focusing becomes too sensitive to the manufacturing tolerances of the valve.

Larger increases in the period can be avoided by inserting a pole-piece into the waveguide or cavity itself, as shown in Figure 5. The axial length of this pole-piece must be less than $0{\cdot}5\ \lambda_g$. This not only helps the matching, but

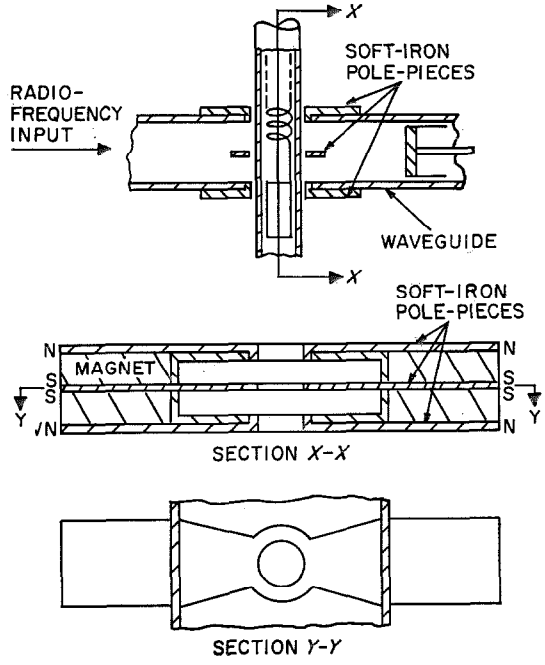


Figure 5—Waveguide-to-helix transducer with an internal pole-piece.

also eliminates the need for two pistons, one above the other. Matching characteristics of a waveguide transducer with and without an internal pole-piece are shown in Figure 6.

1.5 Step Increase in Peak Magnetic Field

The design of a periodic permanent-magnet focusing stack is always based on an unmodulated beam. This is satisfactory unless it is required to focus a power travelling-wave tube. Radio-frequency defocusing forces in the output section may no longer be negligible, and consequently the peak magnetic field in the output section must be increased. A typical increase is about 20 per cent.

Beam profiles in Figure 7 show that a sudden increase in peak field introduces a large beam ripple. A step increase by 20 per cent consisting of at least four steps compresses the beam by about 20 per cent without any apparent change in its profile. Increasing the number of steps beyond four makes no significant improvement. The profiles in Figure 8 correspond to nine steps, the steps increasing either linearly or cubically. There is no practical difference in the final profiles.

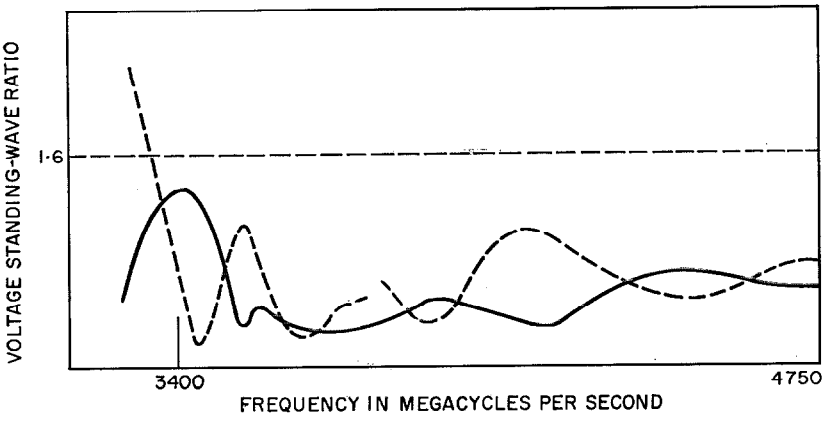


Figure 6—Voltage standing-wave ratio of a waveguide transducer with (broken line) and without (solid line) an internal pole-piece.

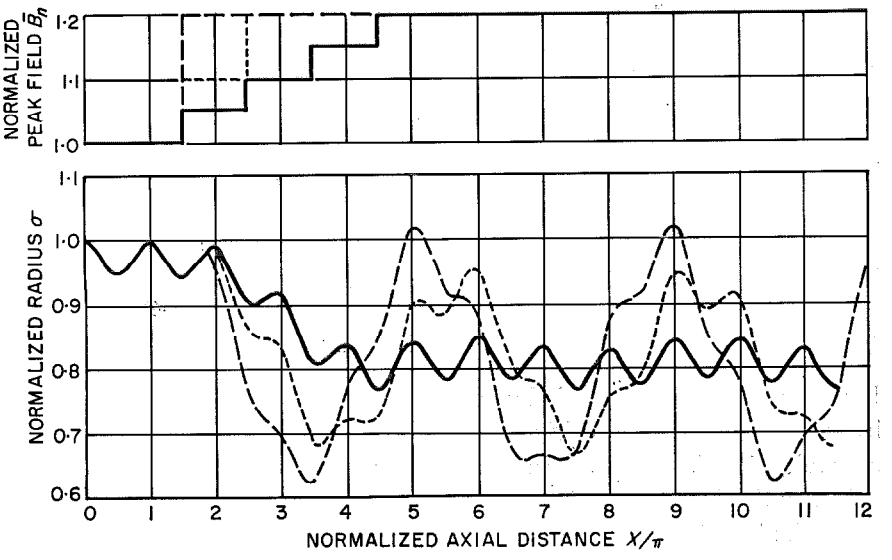


Figure 7—Beam profiles for a step increase in peak magnetic field showing the effect of the number of steps. $K = 0$, $\alpha = 0{\cdot}1$, and $\beta = 0{\cdot}0959$.

1.6 Step Increase in Period and Peak Magnetic Field

Some computations have been done also for a sudden step increase in period alone, and simultaneous increase in period and peak magnetic field. The beam profiles are shown in Figures 9 and 10 respectively. Step doubling of the period in Figure 9 approximately doubles the beam ripple without apparently increasing its maximum diameter.

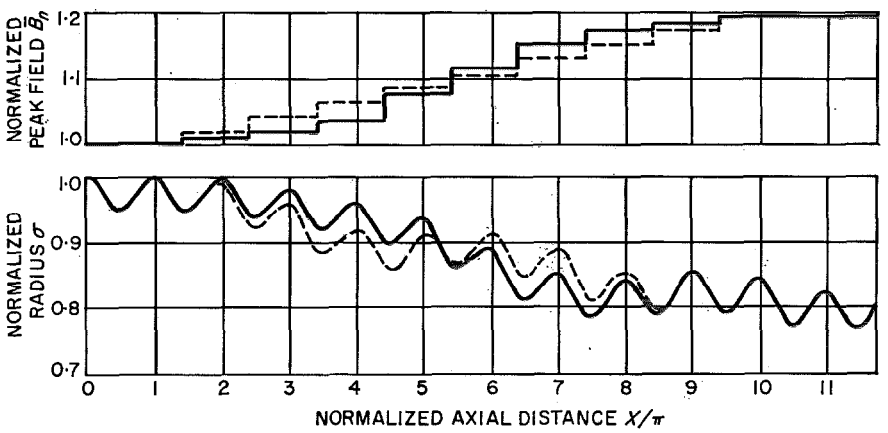


Figure 8—Beam profiles for a step increase in peak magnetic field showing effects of step amplitude. $K = 0$, $\alpha = 0{\cdot}1$, and $\beta = 0{\cdot}0959$.

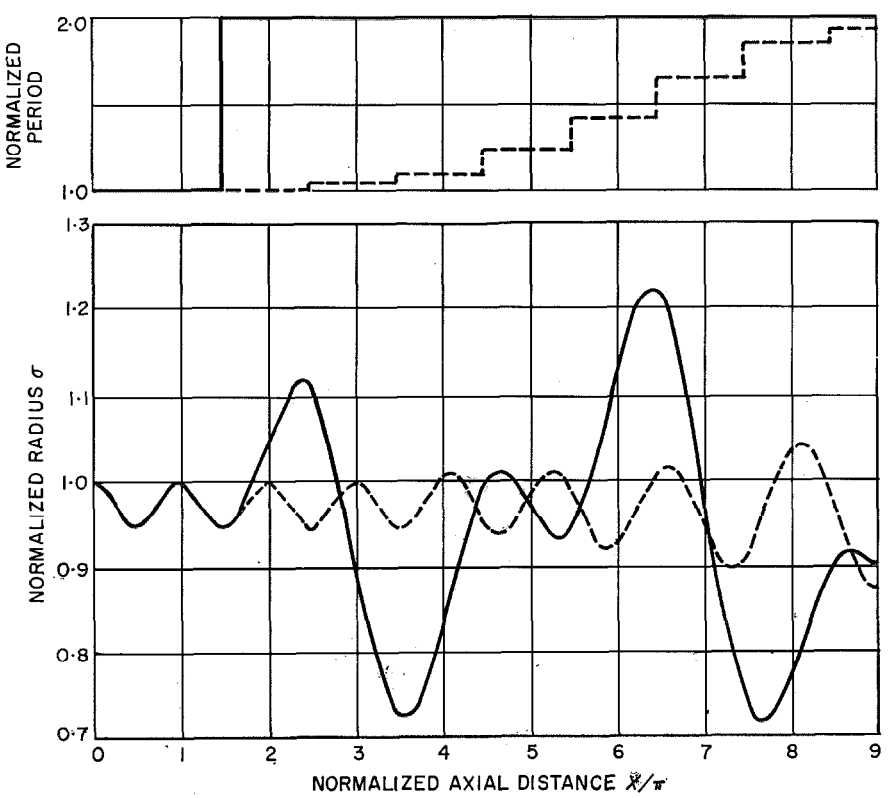


Figure 9—Beam profiles for a step increase in period. $K = 0$, $\alpha = 0{\cdot}1$, and $\beta = 0{\cdot}0959$.

2. Beam Displacement due to Misalignments of Magnetic Cells

Any crookedness of the magnetic field in a periodic permanent-magnet focusing system will deflect the beam from its desired path. We consider here the effect of each section between two pole-pieces, corresponding to a half period, having a magnetic axis parallel to the desired path of the beam, but displaced radially from it.

For the purpose of this analysis, assume that the beam is compressed close to its axis, that is, the beam radius to be smaller than the displacement of the magnetic cell axis.

We use slightly changed cylindrical co-ordinates d, θ, z with the initial entry point and the desired axis of the beam given by $d = 0$; the entry conditions into the nth section are $d = d_n$ and $\theta = \theta_n$. The coordinates of the displaced magnetic axis are D_n and Θ_n, Figure 11.

Considerable simplification in algebra is obtained, and no generality is lost, if θ_n is maintained at zero by rotating the reference co-ordinates at entry to each magnetic cell.

It is shown in the Appendix, Section 6, that if Ω is the angular displacement of the beam axis through one magnetic cell

$$d^2_{n+1} = d_n^2 + 4D_n^2 \sin^2 \Omega/2 + 4D_n d_n \sin \Omega/2 \sin (\Theta_n - \Omega/2). \quad (3)$$

If the different sections are randomly oriented the last term in (3) may be neglected because the average of $\sin (\Theta_n - \Omega/2)$ over all values of argument is zero. Assuming that initially the beam axis coincides with the circuit axis, $d_o = 0$ and (3) reduces to

$$d^2_{n+1} = 4 \sin^2 \Omega/2 \sum_1^n D_n^2 \quad (4)$$

or

$$d_{n+1} = 2n^{1/2} s |\sin (\Omega/2)| \quad (5)$$

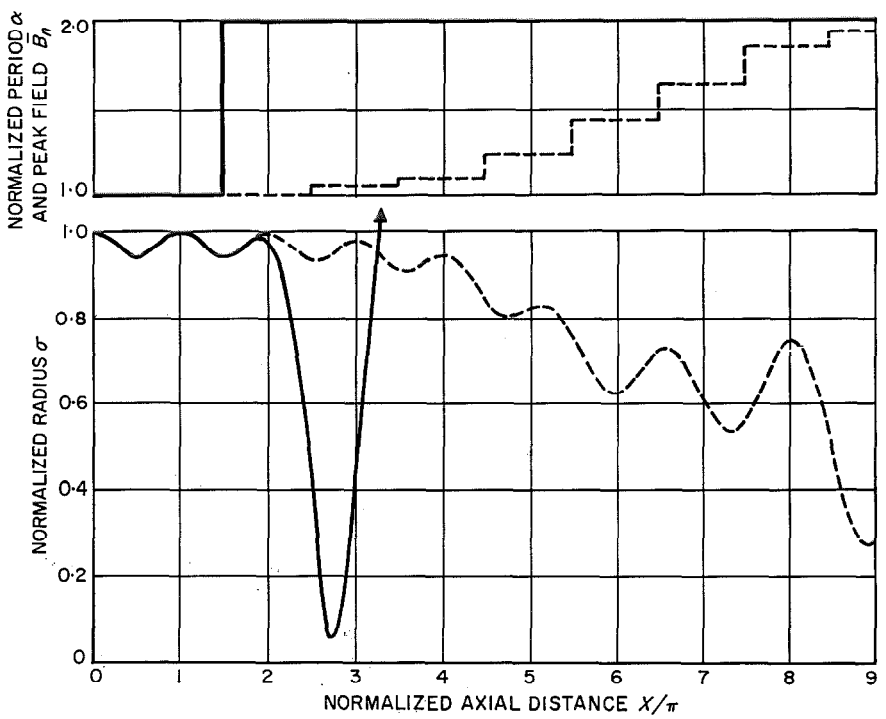


Figure 10—Beam profile for a step increase in period and peak magnetic field. $K = 0$, $\alpha = 0{\cdot}1$, and $\beta = 0{\cdot}959$.

where s is the standard deviation from the intended beam axis.

The angular displacement per magnetic cell Ω is found from Busch's theorem, that is,

$$\Omega = 2(2\alpha)^{\frac{1}{2}}\{(-1)^n - 0{\cdot}5\pi K^{\frac{1}{2}}\} \qquad (6)$$

thus finally, for a shielded cathode

$$d_{n+1} = 2n^{\frac{1}{2}}s|\sin(2\alpha)^{\frac{1}{2}}|. \qquad (7)$$

Can this analysis be applied to all forms of misalignment in periodic circuits? Although formally it does not cover the case where the field is tilted, rather than axially displaced, it would seem that a similar equation would be obtained with perhaps a different definition for the magnetic-axis displacement.

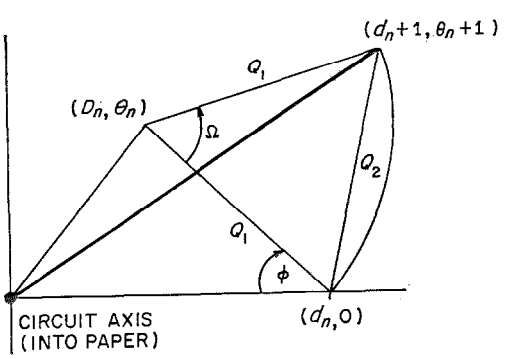
In cases where one or two gross irregularities occur the term $\sin(\Theta_n - \Omega/2)$ in (3) may have to be evaluated.

3. Conclusion

The results show that the undesirable effects of various non-uniformities in the period and magnetic field can be controlled, either by a suitable choice, or by an appropriate variation, of circuit parameters. If α and β are small, as they usually are, the discussed random variations, as well as waveguide gaps, have little effect.

The beam profile for a step increase in the period suggests that one could reduce the number of magnets in a stack by starting with a larger period at the input and tapering down to its optimized value towards the output end of the circuit.

Figure 11—Geometry for electron-beam displacement.



4. Acknowledgments

Thanks are due to Mr. D. Hunter and Mr. M. Lawn of Standard Telecommunication Laboratories computer department for advice on the computer programming.

The approach in Section 2, except for some small modifications, is due to Mr. P. F. C. Burke, and a good deal of this section and the whole of Section 6 have been taken, in changed notation, from his technical note [2].

5. References

1. K. J. Harker, "Periodic Focusing of Beams from Partially Shielded Cathodes," *IRE Transactions on Electron Devices*, volume ED-2, pages 13–19; October 1955.

2. P. F. C. Burke, "Electron Beam Displacement Due to Misalignment in a Periodic Focusing System," Standard Telephones and Cables Internal Technical Note; 1962.

6. Appendix

With reference to Figure 11

$$d^2_{n+1} = d_n^2 + Q_2^2 - 2d_nQ_2\cos\{\phi + \tfrac{1}{2}(\pi - \Omega)\} \qquad (8)$$

$$\frac{D_n}{\sin\phi} = \frac{Q_1}{\sin\Theta_n} = \frac{d_n}{\sin(\phi + \Theta_n)} \qquad (9)$$

$$Q_2 = 2Q_1\sin\Omega/2. \qquad (10)$$

Using (9) to eliminate Q_1

$$Q_2 = 2D_n\frac{\sin\Theta_n}{\sin\phi}\sin\Omega/2 \qquad (11)$$

$$\tan\phi = \sin\Theta_n\{(d_n/D_n) - \cos\Theta\}^{-1}.$$

Substituting in (8)

$$d^2_{n+1} = d_n^2 + 4D_n^2 \sin^2 \Omega/2 \times \{1 + (d_n/D_n)^2 - 2(d_n/D_n) \cos \Theta_n\} + 4D_n d_n \sin \Theta_n \sin \Omega/2 \times \left\{\cos \Omega/2 - \frac{(d_n/D_n) - \cos \Theta_n}{\sin \Theta_n} \sin \Omega/2\right\}$$

$$= d_n^2 + 4D_n^2 \sin^2 \Omega/2 + 4D_n d_n \times \sin \Omega/2 \sin (\Theta_n - \Omega/2). \quad (12)$$

Borivoje Minakovic was born in Zagreb, Yugoslavia, on 16 September 1923. He obtained a B.Sc. (Eng.) degree from London University in 1955.

He worked from 1952 to 1955 for Ferranti Ltd.

Mr. Minakovic in 1955 joined Standard Telephones and Cables and is presently engaged in research and development of travelling-wave tubes.

United States Patents Issued to International Telephone and Telegraph System; May–July 1962

Between 1 May 1962 and 31 July 1962, the United States Patent Office issued 56 patents to the International System. The names of the inventors, company affiliations, subjects, and patent numbers are listed below.

L. J. F. Andre and R. Van Den Bossche, Bell Telephone Manufacturing Company (Antwerp), Mechanism for Extracting a Document out of a Document Carrier, 3 035 726.

R. P. Arthur and N. J. Simkus, ITT Kellogg, Fuse Holder, 3 038 977.

R. W. Blanchard, ITT Laboratories, Transistor Phase-Shift Oscillators, 3 045 191.

F. H. Bray and D. G. Bryan, Standard Telephones and Cables (London), Electrical Calculating Circuits, 3 039 683.

A. E. Brewster, Standard Telecommunication Laboratories (London), Magnetic Information Storage Arrangements, 3 040 304.

A. E. Brewster, Standard Telephones and Cables (London), Magnetic Information Storage Devices, 3 045 216.

A. T. Brown III, Federal Telecommunication Laboratories, Automatic Gain-Control Circuit, 3 036 276.

F. Buchwald and H. Fliegner, Mix & Genest (Stuttgart), Arrangement to Separate Piled Flat Articles From Each Other, 3 035 695.

R. B. Burdett, Federal Telecommunication Laboratories, Multiplanar Printed Circuit, 3 039 177.

J. E. Cox, ITT Kellogg, Two-party Line Discriminating Circuit for Telephone System, 3 046 343.

J. E. Cox, ITT Kellogg, Operator Telephone Circuit, 3 033 940.

B. Dal Bianco and M. Scata, Fabbrica Apparecchiature per Comunicazioni Elettriche (Milan), Multiple Electromagnetic Switch, 3 035 136.

R. C. Ferrar and R. B. Hoffman, ITT Laboratories, Fault Location System, 3 045 113.

H. Fliegner and W. Kastenbein, Mix & Genest (Stuttgart), Equipment to Supply Flat Articles to a Conveying System, 3 039 592.

W. A. Foos, ITT Federal Laboratories, Variable-Bandwidth Intermediate-Frequency Amplifier System, 3 040 268.

W. Fulop, Standard Telecommunication Laboratories (London), Transistors, 3 040 219.

L. B. Haigh and M. G. Kolmes, Federal Telephone and Radio Company, Electric Switch, 3 036 182.

R. J. Heppe, ITT Laboratories, Control Reversal Sensor, 3 044 734.

H. R. Hesse, Federal Telecommunication Laboratories, Collision Course Warning Device, 3 040 314.

G. Hirschfeld, W. Hinz, and H. Fritzsche, Mix & Genest Werke (Stuttgart), Article Sorting Control Apparatus, 3 035 694.

K. G. Hoer, Creed and Company (Croydon) Facsimile Apparatus, 3 044 768.

H. L. Horwitz, G. L. Hasser, J. E. Cox, and S. J. Westhead, ITT Kellogg, Linefinder and Control Circuits, 3 037 084.

R. W. Hughes, ITT Laboratories, Repeater Station for a Bidirectional Communication System, 3 040 130.

R. W. Hutton, A. R. Sigo, and C. J. Adams, ITT Kellogg, Direct-Access Crossbar-Switch Connector System, 3 046 352.

T. M. Jackson and E. A. Sell, Standard Telecommunication Laboratories (London), Cold-Cathode Switching Devices, 3 035 201.

A. J. Katz and R. T. Adams, Federal Telecommunication Laboratories, Frequency-Modulation Radio Altimeter, 3 045 233.

A. M. Klein, ITT Laboratories, Vehicle Suspension and Stabilizing System, 3 035 853.

J. O. Kleinschmidt, Informatic Division of Standard Elektrik Lorenz (Stuttgart), Bistable Trigger Circuit, 3 036 221.

J. B. Lair, ITT Laboratories, Automatic Echo Pulse Recapture Circuit, 3 040 312.

E. J. Leonard, A. Hemel, H. Tarschisch, K. A. Karow, J. C. Gibson, and K. L. Liston, ITT Kellogg, Telephone Connector System Using Controlling Crossbar Switch, 3 036 160.

M. Losher, ITT Laboratories, Servosystem including Quadrature Signal Gate, 3 045 156.

C. Lucanera, F. Kolinsky, and E. J. Annechiarico, ITT Laboratories, Beacon Antenna System, 3 040 319.

H. Lutz, C. Lorenz (Stuttgart), Magnetic Deflecting Yoke for Cathode-Ray Tubes, 3 045 139.

F. P. Mason, Creed and Company (Croydon), Apparatus for Reading Data from a Storage Medium, 3 045 125.

B. McAdams, ITT Laboratories, Multiplex Communication System, 3 040 128.

T. Merritt, Standard Telephones and Cables (London), Lay Plate for Electric Cable, 3 044 244.

B. M. Mindes, ITT Laboratories, Diversity Combining System, 3 045 114.

P. Murden, Standard Telecommunication Laboratories (London), Electromechanical Transducers, 3 036 232.

G. Papp, ITT Federal Laboratories, Wave Detector, 3 041 543.

T. A. Pickering, F. S. Kasper, R. E. Arseneau, and R. F. Maxa, ITT Kellogg, PBX Trunk Hunting in Electronic Switching Telephone Systems, 3 038 969.

W. J. Pohl, Standard Telephones and Cables (London), Electric Discharge Tubes, 3 045 138.

R. Prichard, ITT Kellogg, Coin-Controlled Telephone System, 3 041 398.

L. J. Regis, ITT Federal Laboratories, Astable Multivibrator Pulse Generator Circuit, 3 039 065.

P. Rudnick and R. K. Orthuber, ITT Federal Laboratories, Electron Discharge Device, 3 040 177.

P. C. Sandretto, ITT Laboratories, Navigation System, 3 045 234.

F. P. Smith, ITT Federal Laboratories, Delay-Line Circuitry for Color Television Receivers, 3 042 873.

D. L. Spooner, ITT Federal Laboratories, Voltage Comparator, 3 039 024.

H. F. Sterling and F. J. Raymond, Standard Telecommunication Laboratories (London), Production of Pure Semiconductor Material, 3 044 967.

I. R. Studebaker, ITT Federal Laboratories, Variable-Gain Amplifier, 3 041 545.

F. J. L. Turner and B. S. Mason, Creed and Company (Croydon), Tape Feed Apparatus, 3 044 675.

G. Van Mechelen, Bell Telephone Manufacturing (Antwerp), Associated Circuit for Electrical Comparator, 3 045 186.

R. K. Van Vechten, ITT Laboratories, Tape Transports, 3 035 748.

R. C. Welch, ITT Federal Laboratories, Oscillatory Apparatus, 3 031 878.

R. L. Whittle, C. E. Jones, V. P. Honeiser, and H. S. Margetts, ITT Laboratories, Data-Processing System, 3 036 291.

E. P. G. Wright and J. Rice, Standard Telephones and Cables (London), Electrical Notations Converting Circuits, 3 039 689.

H. Zschekel, Mix & Genest Werke (Stuttgart), Magnetic Storage System, 3 045 213.

Production of Pure Semiconductor Material

3 044 967

H. F. Sterling and F. J. Raymond

This is a process of producing silicon by thermal decomposition of silane using a metal crucible having high thermal conductivity maintained at a temperature lower than the decomposition temperature of silane. A small

charge of silicon is placed in the crucible and heated to the silane decomposition temperature. A steam of pure silane is then directed to the surface of the silicon on the heated surface.

Transistors

3 040 219

W. Fulop

The emitter-base breakdown voltage is increased in a transistor having an emitter and a collector region of one conductivity type and separated by a base region of the opposite conductivity type by providing in the base region a narrow region adjacent to the emitter having a lower impurity concentration than the remainder of the region.

Collision Course Warning Device

3 040 314

H. R. Hesse

This is an automatic device for warning craft operators that they are on a collision course with another craft. The positions and bearings of all aircraft within the range of a plan-position-indicator radar relative to the craft carrying this equipment are recorded on a magnetic storage device. This recorded bearing information is compared with later instantaneous bearing information. An alarm is sounded when the comparison indicates that the craft are on a near-collision course.

Tape Transports

3 035 748

R. K. Van Vechten

This arrangement prevents damage to magnetic tape during drive of the tape in either of two directions. The drive is accomplished by forcing the tape into contact with capstans continuously rotating in opposite directions, by means of respective pinch rollers. The pinch rollers are selectively moved into operative position against the capstans. To prevent tape damage by wrapping around the pinch rollers when the tape is stopped and reversed, these rollers are made with grooves. Protective fingers extend from the mounting plates of the pinch rollers into these grooves. Thus they are not in contact with the tape in normal operation. The pinch roller mounts are also arranged to have a limited axial rotation to assure parallel contact with the capstans.

Facsimile Apparatus

3 044 768

K. G. Hoer

This facsimile system is provided with a magazine for holding a plurality of sheets for successive transfer to a scanning drum. The sheets are separated by rods and the separating members are moved in succession to transfer one sheet at a time to the facsimile scanning drum. Suitable linkages and escapement devices are provided to assure that only one sheet is transferred to the drum for scanning at any one time.

Principal ITT System Products

Telecommunication Equipment and Systems

Automatic telephone and telegraph central office switching systems
Private telephone and telegraph exchanges—PABX and PAX, electromechanical and electronic
Carrier systems: telephone, telegraph, power-line
Long-distance dialing and signaling equipment
Automatic message accounting and ticketing equipment
Switchboards: manual, central office, toll
Telephones: desk, wall, pay-station
Automatic answering and recording equipment
Microwave radio systems: line-of-sight, over-the-horizon
Radio multiplex equipment
Coaxial cable systems
Submarine cable systems, including repeaters
Data-transmission systems
Teleprinters and facsimile equipment

Military/Space Equipment and Systems

Aircraft weapon systems
Missile fuzing, launching, guidance, tracking, recording, and control systems
Missile-range control and instrumentation
Electronic countermeasures
Electronic navigation
Power systems: ground-support, aircraft, spacecraft, missile
Radar
Simulators: missile, aircraft, radar
Ground and environmental test equipment
Programmers, automatic
Infrared detection and guidance equipment
Global and space communication, control, and data systems
Nuclear instrumentation
Antisubmarine warfare systems
System management: worldwide, local

Industrial/Commercial Equipment and Systems

Distance-measuring and bearing systems: Tacan, DMET, Vortac, Loran
Instrument Landing Systems (ILS)
Air-traffic control systems
Direction finders: aircraft and marine
Ground and airborne communication
Data-link systems
Inverters: static, high-power
Power-supply systems
Altimeters
Flight systems
Railway and power control and signaling systems
Information-processing and document-handling systems
Analog-digital converters
Mail-handling systems
Pneumatic tube systems
Broadcast transmitters: AM, FM, TV
Studio equipment
Point-to-point radio communication
Marine radio
Mobile communication: air, ground, marine, portable
Closed-circuit television: industrial, aircraft, and nuclear radiation
Slow-scan television
Instruments: test, measuring
Oscilloscopes: large-screen, bar-graph
Vibration test equipment
Magnetic amplifiers and systems
Alarm and signaling systems
Telemetering
Intercommunication, paging, and public-address systems

Consumer Products

Television and radio receivers
High-fidelity phonographs and equipment
Tape recorders
Microphones and loudspeakers
Refrigerators, freezers
Air conditioners
Hearing aids
Incandescent lamps
Home intercommunication equipment
Electrical housewares

Cable and Wire Products

Multiconductor telephone cable
Telephone wire: bridle, distribution, drop
Switchboard and terminating cable
Telephone cords
Submarine cable
Coaxial cable, air and solid dielectric
Waveguides
Aircraft cable
Power cable
Domestic cord sets
Fuses and wiring devices
Wire, general-purpose

Components and Materials

Power rectifiers: selenium, silicon
Parametric amplifiers
Transistors
Diodes: tunnel, zener, parametric
Semiconductor materials: selenium, germanium, silicon
Capacitors: wet, dry, ceramic
Ferrites
Tubes: power, transmitting, traveling-wave, rectifier, receiving, thyratron
Picture tubes
Relays and switches: telephone, industrial
Magnetic counters
Resistors
Varistors
Fluorescent starters
Transformers
Quartz crystals
Crystal filters
Printed circuits
Hermetic seals
Magnetic cores

VOLUME 38 • NUMBER 3 • 1963